DOCUMENT RESUME

ED 129 131 HE 008 140

AUTHOR Dolan, W. Patrick
TITLE The Ranking Game: The Power of the Academic Elite.
INSTITUTION Study Commission on Undergraduate Education and the Education of Teachers, Lincoln, Nebr.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE 76
NOTE 121p.
AVAILABLE FROM Study Commission on Undergraduate Education and the Education of Teachers, University of Nebraska - Lincoln, 338 Andrews Hall, Lincoln, Nebraska 68508 ($2.00)

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS Educational Quality; *Educational Research; *Educational Status Comparison; *Evaluation; *Faculty Evaluation; Graduate Students; *Graduate Study; *Higher Education; Professional Education; School Surveys; Universities

ABSTRACT

This book reviews the Cartter and Roose Anderson reports on graduate education published by the American Council on Education (ACE) in 1966 and 1970, respectively. The author sets out three questions that require attention in looking at a process such as the ACE rankings: (1) who is evaluating whom? (2) what is being evaluated? and (3) for what purpose? A broad outline of the present context of judging educational goals and behavior is presented. The criteria for judging "quality of graduate faculty," described as scholarly competence and achievements, is examined. The studies themselves are examined, including their assumptions and methodology against this background as well as their rhetoric as it effects their claims and use. Some of the uses of the ACE rankings, within and without institutions of higher education, are discussed along with some alternatives to these two studies and their uses. (JMF)

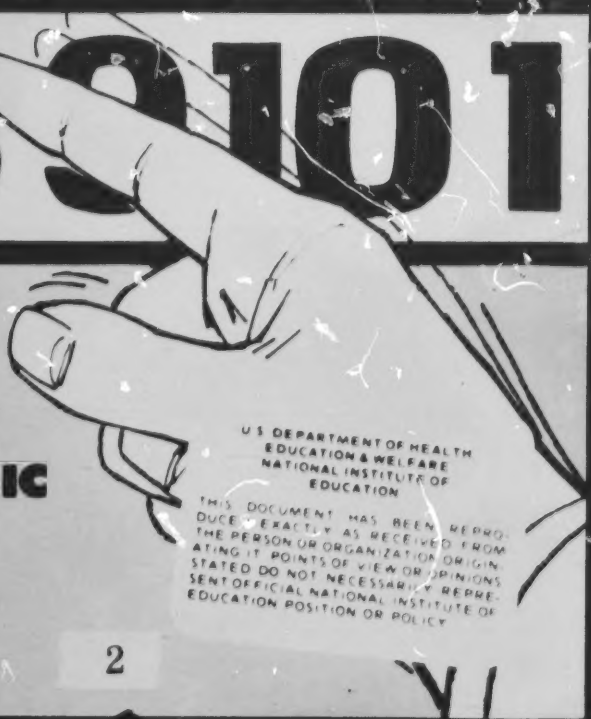
****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

THE RANKING GAME

by
W. PATRICK DOLAN
preface by
WILLIAM A. ARROWSMITH

ED129131

1 2 3 4 5 6
7 8 9 10 1



THE POWER OF THE ACADEMIC ELITE

HE008140

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE RANKING GAME:

THE POWER OF THE ACADEMIC ELITE

By W. Patrick Dolan

Preface by William A. Arrowsmith

Evaluation of Higher Education Committee
of the Study Commission on Undergraduate Education
and the Education of Teachers

1976

3

This monograph prepared by the Evaluation of Higher Education Committee of the Study Commission on Undergraduate Education and the Education of Teachers (William Arrowsmith and W. Patrick Dolan, chairmen) is one of a series of Study Commission publications and does not represent an official position of the Study Commission. The book is a study document for distribution to those associated with the work of the Commission. Requests for this book and other Study Commission publications should be addressed to: The Nebraska Curriculum Development Center, Andrews Hall, University of Nebraska, Lincoln, Nebraska 68588.

Publication of this document at the University of Nebraska Printing and Duplicating Service was funded with a grant from the U.S. Office of Education, Department of Health, Education and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

OF EXACTITUDE IN SCIENCE

. . . In that Empire, the craft of Cartography attained such Perfection that the Map of a Single province covered the space of an entire City, and the Map of the Empire itself an entire Province. In the course of Time, these Extensive maps were found somehow wanting, and so the College of Cartographers evolved a Map of the Empire that was of the same Scale as the Empire and that coincided with it point for point. Less attentive to the Study of Cartography, succeeding Generations came to judge a map of such Magnitude cumbersome, and, now without Irreverence, they abandoned it to the Rigors of sun and Rain. In the western Deserts, tattered Fragments of the Map are still to be found, Sheltering an occasional Beast or beggar; in the whole Nation, no other relic is left of the Discipline of Geography.

From *Travels of Praiseworthy Men* (1658), attributed to J.A. Suarez Miranda

Quoted in: *A Universal History of Infamy* by Jorge Luis Borges (Dutton: 1972), p. 139.

TABLE OF CONTENTS

Page

PREFACE *by William Arrowsmith* vii

INTRODUCTION *by W. Patrick Dolan* xv

ACKNOWLEDGMENTS xvii

CHAPTERS

I. The Context of the Rankings 1

II. The Academic Profession 11

III. The Study Itself: Assumptions and Rhetoric 25

IV. The Study Continued: Methodology and Content 39

V. The Use and Abuse of Statistical Hearsay 61

VI. Summary 73

APPENDIX A–The Welch Report on Graduate Education in Religion 83

APPENDIX B–The Rank and File 93

BIBLIOGRAPHY 107

PREFACE

by William Arrowsmith

Patrick Dolan's study of the American scoring system in higher education touches on issues much more important and revealing than the absurdity of statistical efforts to quantify quality or otherwise square the circle.

What Mr. Dolan supplies—to those willing to read between and beyond the lines—is what, to my mind, is usually missing from the common accounts of malaise and academic "mindlessness" (to borrow Charles Silberman's all too happy unhappy phrase): a sense of the pervasiveness and range of academic and bureaucratic "bad faith," its elegantly interlocking mechanisms, the prodigious beauty of its complicating organization; and the nice ramification of its effects and causes through institutions of higher learning to the national and even supra-national grids that rationalize and reinforce the institutional guilds. The function of this massive apparatus is to protect professions and scholars (but also bureaucrats and politicians)in what they do and do not do, but above all from the jeopardy of imaginative vision and action and the risk of real freedom. In engineering jargon "bad faith" is a form of systemic dysfunction called "suboptimization," by which is meant a high degree of efficiency in the production of an undesirable or irrelevant output. Its cause is almost always deliberate or careless inattention to any ends which are not simply marginal or proximate.

Suppose, for the sake of illustration, that an intelligent dean, perhaps an educator to boot, asks a chairman "to improve his department's performance." The dean, we may suppose, is exercised about some intangible called "excellence." The chairman and his senior colleagues, faced with this exhilaratingly vague challenge, respond suboptimally—that is, in ways they see normalized everywhere about them in the institution. They respond, that is, by mindlessly (or cunningly) tightening up what they like to call "standards": the syllabus is modified slightly, a few missing subspecialties are deftly stitched in, dubious junior colleagues and "unstable" graduate students are

sent packing; there is a flurry of memoranda on some timid "new" departure, memos on the best means of achieving greater "professional visibility" are circulated, and so on. The final illusion created is that of a taut little ship, a sound professional hand on the tiller, sprucely sailing along with the favoring tradewinds currently blowing from whatever fashionable quarter, making briskly for nowhere.

Nowhere. In American universities nowhere is not an unknown destination. It is, in fact, a familiar port-of-call. It is a number–whose "worth's unknown although his height be taken." The number sought is a higher grade on the list of the top twenty graduate departments, all rank-ordered, discipline by discipline, in those lustral documents known as The Cartter Report and the Roose-Anderson Report (here examined closely by Mr. Dolan in this study). Published under the auspices of The American Council on Education, these documents were originally commissioned by three federal funding agencies: the National Academy of Sciences, the National Institute of Health, and the Office of Education. They purport to represent "expert opinion"–that is, the judgment of the professoriate on its own performance. The rankings doubtless indicate *something*, but as an index of quality they are of extremely dubious value; in fact, they are little more than quantified gossip or hearsay. Nonetheless, for obvious reasons, their appeal is enormous simply because in a milieu devoid of final goals, they suggest an available target: the illusion of an end, or achievable virtuosity in proximate ends. To unimaginative or stupid administrators they are miraculously appealing simply because they assign a hard number (to *two* decimal places!) to quality and indicate clear models of imitable virtue and respectability. For clearly a high ranking is as good as gold; indeed, it *is* gold. Even a slow-witted administrator can grasp the import of a high ranking on a rating list funded by those agencies charged with the disbursement of public funds and accountable to unpredictable and perhaps irritable congressional committees.

Even administrators, that is, can grasp, beneath all the plausible talk about the wisdom of having peers judge peers and the incontrovertibility of "expert" opinion, the profile of still another form of institutionalized and bureaucratic "bad faith." The responsibility for assigning public funds is a risky business at times; mistakes–*ascertainable* mistakes–can be costly for a man's career. But the responsibility can be avoided by the simple device of summoning "expert opinion" to testify (and since, in the case of the Cartter Report and the Roose-Anderson Report, the experts are all anonymous, they

cannot be required to testify in person or explain their opinions). A congressman angrily demanding to know why *his* district, *his* university, has been consistently scanted by a funding agency can be abruptly silenced or shamed by this published consensus of anonymous experts. The rankings of course are most warmly defended by those institutions which have been most favorably ranked (and vice versa)—and which, in fact, supply the bulk of the respondents.

But the overall *effect* of these reports seems quite clear. They tend, first, to make the rich richer and the poor poorer; second, the example of the highly ranked clearly imposes constraints on those institutions lower down the scale (the "Hertz-Avis" effect). And the effect of such constraints is to reduce diversity, to reward conformity or respectability, to penalize genuine experiment or risk. There is, also, I believe, an obvious tendency to promote the prevalence of disciplinary dogma and orthodoxy. All of this might be tolerable if the reports were tolerably accurate and judicious, if they were less *prescriptive* and more *descriptive;* if they did not pretend to "objectivity" and if the very fact of ranking were not pernicious and invidious; if they genuinely promoted a meaningful "meritocracy" (instead of simply perpetuating the *status quo ante* and an establishment mentality). But this is precisely what they cannot claim to be or do.

Patrick Dolan's study of the Cartter Report and the Roose-Anderson Report makes it unmistakably clear to anyone except the most naive partisan of the *status quo* that these putatively "statistical" assessments of opinion about quality are worthless, moral and intellectual absurdities. In their veiled prescriptiveness, their aura of hairline numerical accuracy, and their implicit advocacy of a single educational norm, they are patently tendentious and regressive documents. Nonetheless, they are enormously influential, probably because they seem to rationalize and mythologize the *fact* of Riesman's "academic revolution"—because they describe, with all the weight of "expertise" and statistics in a society distrustful of "quality," the accelerating trend toward uniformity and conformity visible at every level of American education. They are also influential because they are concerned with graduate training, which, as everybody knows is the apex of a pyramid—an apex which, in its angular definition, geometrically determines the base—secondary and undergraduate education—which supports it and is pointed toward it. Academic prestige and respectability, federal and foundation funding, the ability to attract intelligent students, the content of disciplinary and institutional

programs, professional priorities, individual goals and institutional behavior—all are intimately and profoundly affected by the documentary inanity of these rankings.

In themselves absurd and often patronized by those who profit from them most, they are, in *actual* educational and political use, a real influence on the behavior of those who recognize their absurdity; their influence upon those who do not recognize how factitious and spurious and meaningless these statistics are—that is, trustees, the general public, laymen, alumni, students, and legislators—is extremely potent for good or for ill. Ultimately, persistently, these documents have a tangible influence on the life of the mind (increasingly synonymous in our society with colleges and universities)—on revenues, programs, priorities, recruitment, even institutional stability itself. And that influence is, I believe, predominantly negative. For while it may be of some national significance that mediocrity should have imitable models of excellence, there is immense danger in the conformity that comes of such imitation. The life of the mind has never flourished when constraints are placed upon the nature, scope, or limits of inquiry; when a formulative Mandarinism and deference to the tried-and-true replaces vivid ... doxy, the leap of the imagination into the disciplinary unknown.

And *these* constraints are precisely the price of the "normative" model built into the A.C.E. Reports. The academic world is not, in its present state, either notably open-minded or temperamentally disposed to risk. These documents, while pretending to promote "standards," are in fact formidable professional sanctions against daring, diversity, openness. And, in a time of academic recession, the constraints imposed may become exceptionally daunting. When graduate programs which fail to "pull their weight" (which can mean either by attracting students or commanding a high ranking in the A.C.E. Reports) are everywhere being phased out or cut back, high-risk programs (e.g. interdisciplinary studies in particular or any "approach" not sanctified by mainstream disciplinary practice) can rarely even succeed in obtaining institutional approval, let alone the "outside" funding which might reduce the "inside" opposition. The very tight "grid" of vested, local, and professional interests (embodied in their functional modular unit, the department) tends to discourage significant—as opposed to trivial—innovation. And this conservative grid is reinforced from above by an extremely complex and powerful national "grid" which represents the same forces as those bonded locally in the department.

Mr. Dolan vividly describes the anomaly of the Educational Testing Service, a non-accountable private institution (representing, again, the views of the university professoriate), which effectively controls access by students to higher education, just as the A.C.E. Reports are used to define graduate performance and to screen out threatening alternatives. Here we have a good example of two mutually reinforcing institutions, both working along what can only by courtesy be called decently suboptimal lines, intersecting so as to "certify" schools and students to each other in the interests of the single-model "research university" and its faculty. To these institutions must be added a host of others—the Association of Graduate Schools, the A.C.E. itself (higher education's "lobby" in Washington), the American Council of Learned Societies, the syndicalism of the A.A.U.P. (which now vigorously champions, in the name of "academic freedom," what is elsewhere called "seniority," "job-security" and even "feather-bedding"), as well as the institutions which effectively fund much of graduate education, i.e. the foundations, the state and federal bureaucracy, etc. The formidable "mesh" of all these various institutions, almost invariably committed to the single-model university, constitute a force which, both locally and nationally, have effectively made significant academic reform virtually impossible. A vigorous administrator may, it is true, produce significant local change, but no sooner is his back turned or his attention diverted, than local and national grids help to recreate the *status quo ante*. And the reason is painfully obvious, as Mr. Dolan's frightening account of the genesis of the A.C.E. Reports makes clear. Cumulative and collective "bad faith," ruthlessly and mindlessly (there is no ruthlessness like the absent-mindedness of those who have never even imagined that their benevolence could be questioned or faulted) intersecting with the "bad faith" of the funders in the bureaucracy and the foundations, combing to deny even the *idea* of optimal goals.

Demonologies, however tempting, are out of order. Bad faith, after all, is not conspiracy or even sin, it is simply a refusal of one's true freedom, a form of privileged and even contrived "irresponsibility." The man of bad faith disowns the fact of his own power, which is to say his responsibility; and the disavowal is politically and educationally fatal. Collective "bad faith" is merely the conformity of the irresponsibles, men of presumptive "good will" combining, through habit and role and the fear of risk, in inaction or wrong action masked by anonymity. A man who is "just doing his job" is to that very degree usually suboptimizing, refusing his freedom or using it in deliberately or unconsciously marginal ways. Bureaucracies are by tendency

suboptimal; not only are they designed to "implement" policies and priorities decided elsewhere, but they are designed to be efficient and therefore deprived of any true sense of a larger or general purpose. And they are notoriously subject to that insensitivity which is born of routine and limited goals.

What is frightening about Mr. Dolan's scenario of the genesis of the A.C.E. Reports is the utter absence from it of any serious *educational* purpose; at no point, it would seem, did any generous or optimal educational view prevail. The sense of an *overview*—an ample, complex sense of the problems involved, their ramifications and implications—is stunningly absent; there is no sign of genuine leadership anywhere, merely "advocacy" of diverse points of view whose inter-relatedness would seem to have been an "act of faith" in the virtue of "ongoing" educational or bureaucratic activities. In a country capable of going to war in a fit of absent-mindedness, it is perhaps not surprising that lobbyists and bureaucrats should tend to act as though the concerting of their different purposes was *ipso facto* beneficial. But it is, to say the least, disappointing that those who represent "higher education" and who "formulate educational policy" should exhibit such utter blindness to their own assumptions and activities. To the bureaucrats of the National Science Foundation it would no doubt have been *useful* to possess an "exactly" rank-ordered hierarchy of scientific talent, discipline by discipline. And, as Mr. Dolan's account makes clear, the officials of the A.C.E. predictably perceived the N.S.F. project as a threat to the professoriate, and rose in defence of their own constituency and the principle that only experts should judge experts. Merely bureaucratically, the whole project made superb sense. By sharing costs, responsibility (and risk) could be shared; and the bureaucrats could be spared political embarrassment by shifting responsibility to the academic "experts," who in turn, in the name of "impartiality" and "objectivity," were to be protected by complete anonymity. The result was a wholly "impersonal" document, a triumph of contrived "objectivity" which, by means of technical statistical jargon and anonymous "expert" opinion, effectively removed the obscenity of "value" and "value-judgments" from the funding-process. *Nobody*, after all, was responsible for the rankings except the anonymous experts; *everybody* was protected from the risk of responsibility. Finally, even the possibility of ugly and jolting surprises was removed by the simple device of having "deans of leading universities" nominate the respondents in the various disciplines. Since there were no ugly surprises, the Reports could in fact be presented as a plausible vindication of what every-

body already "knew" anyway. The Reports are in this sense contrived predictions—whose contrivance—in its obviousness—appears not to have been wholly evident to its designers. But the contrivances are utterly apparent in (1) the assumptions (single-model excellence); (2) the selection of respondents, and (3) the questions (which, as Mr. Dolan points out, ignored teaching, departmental milieu, and all the crucial intangibles which go to make up a genuine intellectual community).

Mr. Dolan's scenario, then, is the familiar tragedy of suboptimal determination of national and cultural goals. We are used perhaps to the fact that government agencies pursue, frequently, wholly unrelated and even inconsistent ends; and we are even used to the absence of political leadership that makes such bureaucratic behavior possible. But education defines after all our intentions about the future of the species, or that part of the future which is our own children. Education defines, or should define, our sense of what is required to create and preserve our own humanity. One can perhaps afford to be suboptimal about the fate of the coyote or the pinon pine or oil-shale reserves (there is in fact little likelihood of being anyting else). But education cannot be decently abandoned to the bureaucrats or the *mauvaise foi* of the academic profession and its lobbyists. We are, for better or worse, a "knowledge-society," and our fates will be defined, for better or worse, in the schools and universities. The life of the mind requires, like culture, like education, not homogenized programs, learning, curricula, colleges, and communities, but precisely diversity, the *possibility* of risk, the *necessity* of abundant "otherness." The national tide, at least in education, is still running strongly toward centralization, to the mindless annulment of diversity and pluralism. And the A.C.E. Reports are symptomatic of this current and the deliberately fostered tendency of the whole mindlessly purposive national grid. The Cartter Report and the Roose-Anderson Reports are monstrosities, not simply because they are the patent product of reflex bad faith and suboptimization, but because they are the regressive instruments of *standardization* which has no valid cultural or human purpose; which serves merely governmental, professional, administrative, or bureaucratic convenience. They simply have no educational purpose which is compatible with the needs of our culture or with the life of the mind.

Mr. Dolan suggests that, if we must have such reports, they should be *descriptive*, not subtly but unmistakably *prescriptive*. And this seems to me unquestionably right. But the point goes, I think, far beyond our need for

educational equity, valuable as that may be. The greatest single need in the country is, I would argue, for diverse communities, communities of the mind or spirit, above all, which is what the ideal college or university should be. Such communities are extremely difficult to build, extremely fragile by their very nature. (I doubt that more than a dozen true intellectual communities have ever existed on this continent—but the history of the country is inextricably involved with those dozen or so communities.) At present *all* true culture, whether "minority" or "academic" or "regional" is threatened by the enormous power of what Sapir has called the "Civilization"—all the devices of progress and sophistication which came into being in order to propagate and enable the Culture, but which have now largely, if not totally, usurped it. The national "grid" of education, devices like the Cartter Report (or the airlines, big labor, big business, etc.) are in fact the instruments of the Civilization, and they are now *wholly* antithetical to the growth of indigenous, coherent, "cultures" everywhere—from tribal Islam to the Hopis to the desperate effort to create true intellectual or academic communities. Everywhere we look we see in fact the synthetic elitism of the Civilization usurping or destroying what is left of genuine Culture, and apparently, in its suboptimal dynamics, irresistible and perhaps irreversible. The danger, the crucial danger, is that of failing to recognize the difference, of bowing to the great organized engines of the Civilization as though they still contained the old spiritual imperatives of education and culture—or the education and culture which must be saved from "meritocracy" and synthetic elitism.

INTRODUCTION

When the American Council on Education's ranking of graduate departments appeared in 1966, they occasioned immediate debate, and one of the first voices raised in protest was that of William Arrowsmith. Since his article "The Shame of the Graduate Schools," in *Harpers* in 1964, he had been an outspoken critic of the single model prestige university and its unquestioned endorsement by the academic profession. The rankings of graduate schools by the American Council on Education in 1966 and again in 1970 appeared to him to be perfect examples of this sickness, and a symptom of things to come. As this book, and especially its appendices attest, his fears were more than justified.

By the time the second ACE ranking had appeared in 1970, Arrowsmith had found some strong support for his criticism within the U.S. Office of Education in the person of Donald Bigelow. Throughout the 1960's Bigelow had been giving impetus to federal programs attempting to change the manner in which teachers were educated and trained. In the process there had been a great deal of experience working with institutions of higher education, including both schools of education and liberal arts faculty. The intransigence of the structures, the power of the professional "ideal," and the often unexamined assumptions of higher education had all been enountered and documented in those experiences. Bigelow felt the time was right for a careful examination of these structures and assumptions, and agreed that Arrowsmith's direct confrontation of the ACE rankings and their ramifications was the best approach.

At that time, in 1971, Donald Bigelow was in the process of forming a national study commission to examine undergraduate education and the education of teachers, which Paul Olson of the University of Nebraska was to direct. That commission's concern was with the aids and hinderances to real diversity, both cultural and educational, in this country's schools and colleges. A Committee of the Evaluation of Higher Education was formed as a

part of that larger commission, and William Arrowsmith and I became co-directors. We have worked closely together on this project since 1972, and although I have been the one to finalize this essay, his presence has been continual and indispensable.

The ideas and concerns presented here, then, began in the aftermath of the ACE rankings and grew into broader questions as each of us brought our own perspectives and values to this project. At the same time this essay represents my own statement about the present art, science, and politics of evaluation of higher education. The ranking of graduate departments by the American Council on Education is, I fear, not a naive aberration, but a bellwether. We have already begun to see a proliferation of these kinds of rankings, but the original provides an excellent entrance into higher education in the United States today. It is built upon judgments of quality, implicit and explicit assumptions about the structure and function of the university, and political and professional practices, which I wish to examine and question here. Finally, for all the statistical accoutrements and quasi-scientific language, it is a political document, and I will treat it as such.

W. Patrick Dolan

Rancho Santa Fe, California
September, 1975

ACKNOWLEDGMENTS

I wish to express my gratitude to William Arrowsmith, Donald Bigelow, and Paul Olson. Each, with his support and encouragement, helped make this book possible. I would also like to thank Malcolm Shaw for his generosity during the final writing of the manuscript. Both Charles Andersen, then at the American Council on Education, and Alan Cartter, who was Chancellor of New York University at the time, were generous and candid in personal interviews. Margaret Shearon helped a great deal in the early research, and Jan Pieper and Jim Bowman of the Study Commission staff shepherded the whole process with more than ordinary patience and understanding. Finally, to the many respondents to our letters, some of whom appear in the text, many of whom do not, William Arrowsmith and I offer a special thanks.

I. THE CONTEXT OF THE RANKINGS

The Central Questions

In 1966, Allan Cartter, then Vice-President of the American Council on Education, published under the council's aegis a ranking of the graduate schools, department by department, entitled: *An Assessment of Quality in Graduate Education.*[1] It was a study which addressed itself to the question of quality in higher education at the graduate level, and attempted to assess and rank that quality based on institutional reputation among some four thousand scholars. The study was also expressly designed "to examine critically the available techniques of evaluation." At that time Cartter promised a similar study not later than 1970, in order to avoid "freezing" the reputations and rankings that appeared in his study. Dutifully the American Council on Education (ACE) published a similar report in 1970, this time compiled by Kenneth D. Roose and Charles J. Andersen, with the somewhat more modest title: *A Rating of Graduate Programs.*[2]

Appearing as these two studies did, in the almost total absence of any other kinds of evaluation available to the public, and wielding the power and prestige of the American Council on Education, they became extremely important documents. They are presented in both cases as a set of ratings which could and should, in the authors' minds, be used as a basis for national policy and funding decisions, as well as for institutional and graduate student use. We shall look later at the presence of their judgments in practically all the popular guides to graduate schools, but, more importantly, there is every indication that these studies have provided a model and impetus for additional

---

[1] Allan M. Cartter, *An Assessment of Quality in Graduate Education* (Washington, D.C.: ACE, 1966). This study will be referred to later in this report as Cartter, *An Assessment.*

[2] R.D. Roose and C.J. Andersen, *A Rating of Graduate Programs* (Washington, D.C.: ACE, 1970). This study will be referred to later in this report as Roose-Andersen, *A Rating.*

rankings in the name of better public knowledge and better planning in higher education.[3]

There seems to me to be three questions that require some attention in looking at a process such as these rankings: *Who is evaluating whom? What is being evaluated? For what purposes?* The answers to these questions should provide some explanation as to why certain assumptions were made, certain questions were asked, and why rankings were the result, instead of some other kind of assessment. Since these rankings begin so cleanly, as if we all knew and agreed upon what quality education is, they represent an excellent example of the unstated goals and prevailing and assumed criteria for judgment now operating in higher education. And since *what* is being evaluated is almost totally undefined in the word "quality," it would be helpful to begin with a broad outline of the present context of judging educational goals and behavior—in short, the state of the art of assessment of higher education at present. This chapter will attempt to do this.

Further, in each of the two studies and their questionnaires, the central question concerning "quality of the graduate faculty" is qualified with the single instruction: "Consider only the scholarly competence and achievements of the present faculty." It would appear necessary, then, to look, in addition, at the academic profession's present ideas on "scholarly competence," to understand what the sources of a judgment about this area might be, and what effect this as the sole qualifier might have on the resulting responses. The second chapter addresses these questions.

The third and fourth chapters look at the studies themselves—their assumptions and methodology against the background of the first two chapters, and their rhetoric (statistical?) as it effects their claims and use. The reports are of high statistical quality internally, especially the first one by Cartter, and a great deal of time and effort was spent in this regard. Yet little has been said about the external validity of "reputational studies" in such a context as this, especially with a close look at the universe of correspondents defined by the studies themselves.

---

[3]See Appendix A and B for examples of the Welch Report and other evidence that rankings are in full swing.

The movement from vague questions of undefined "quality" to highly quantified results has led to a constant tension throughout. The authors always admit that the rankings are only "a survey of informed opinion," and yet push out analysis to two decimal places, and test the validity of such methodology to the point where an impression is created of the presence of a fine-grained scientific analysis. Then one realizes that this is page after page of statistical correlations about judgments based on hearsay evidence about undefined qualities and undetermined goals. It is that movement from an interesting opinion poll to the pretense of precise rankings and a formidable quantified instrument that is most important to capture and document. For this is the most subtle and misleading transition in the studies.

Uses of such studies are of course out of the hands of those who undertake them. Yet certain forms of evaluation are more suited to certain kinds of uses and, in fact, often determine that use. Numerical rankings, for instance, can be applied in much different ways than qualitative description and by their nature almost guarantee such use. The fifth chapter examines some of the uses of the ACE rankings within and without institutions of higher education, and the sixth talks about what alternative paths lie open.

How such quasi-information is used, of course, depends a great deal on the environment in higher education described in the first two chapters. If there were a great deal of solid comparative data publicly accessible, that would make the appearance of documented hearsay one thing. Unfortunately, such is not the case. If there were a great deal of clarity about the components of quality in higher education, then a vague questionnaire would be nothing more than an opinion poll. But, such is not the case. If the powerful mechanisms that control higher education were publicly accessible and responsive, then perhaps a variety of institutions and openness to different kinds of "quality" would be imaginable. But a profession in the process of closing its ranks, and a series of private interlocking associations, move toward other goals. The dynamics of power in higher education and the level and impact of knowledge about performance has direct bearing on how and why the ACE rankings took the form and shape they did. An examination of that process must be the first question.

The Universities

If anything became clear out of the 1960 literature about change in

higher education, it was that the American university was, in Donald Schon's words, a "dynamically conservative" institution.[4] The forces operating within and without to preserve it as it was, and to continue its own self-contained functions, were formidable. The universities could and did take a great deal of pressure from without, during those years, and did what was necessary to come to terms with that pressure. There was some interdisciplinary tinkering, a few colleges within colleges for a small group of students, some changes in grading, and so forth, but the basic structures seem to have undergone little if any significant change. That was no mean feat accomplished by an institution that at first glance appeared so vulnerable. But the longer one compares the American university with other institutions in this society, public or private, the more one is impressed with the sameness and the invulnerability of its structures.

The research and resultant literature has been hard put to explain the unyielding nature of the structures of higher education and their number. The litany is endless: their single way of measuring units of learning by credit hours, constant use of the department as the modular academic unit, the fixed apparatus of professorial rank and accompanying promotional procedures, acceptance of the classroom and lecture method as the only learning experience certifiable, the great percentage of the work assigned to students only on an individual basis, even the growing sameness of the content of the courses in each discipline, the use of academic majors, the same admission criteria and formula for all kinds and shapes of colleges and universities, and on and on it could go.

But the most puzzling aspect of all is the reluctance to question or tamper in the slightest way with this series of formulas for higher education.

---

[4]This phrase is used by Schon to describe the enormous output of energy by institutions to preserve and hold stable the patterns and structures of the ongoing system. His analysis of the relationship between institution and individual has interesting applications in the case of the university and its professionals. For example, "The power of social systems over individuals becomes understandable only if we see that social systems provide for their members not only sources of livelihood, protection against outside threat and the promise of economic security, but a framework of theory, values and related technology which enables individuals to make sense of their lives. Threats to the social system threaten this framework" (Donald Schon, *Beyond the Stable State*, Random House, 1971, p. 51.)

The unyielding nature of the formula seems to rest on some surety that this is the perfect way of teaching and learning. It proceeds as if there were no ambiguity as to higher education goals, that unanimity existed around correct methodology, and that there was total agreement about the structures to carry these off. Yet these are areas of constant debate. It is this puzzling contradiction that highlights the dilemma.

Every administrator or teacher in an individual institution who has attempted to move some tiny particle of the structure is always amazed at the strength of the entire fabric. They seek the causes of the rigidity *within* a particular institution where they find themselves. But the answer may be that the causes of structural rigidity in American higher education do not exist within any individual institution. Higher education in the United States has been progressively surrounded with superordinate structures such as accreditation, certification, the testing-entrance procedures, the distribution system for monies, the growing strength and values of the academic profession itself. These may represent the real source of intransigence and the difficulty of effecting change at any given institution.

Accreditation

The process of becoming an acceptable institution of higher education in the United States is a model for the other accompanying structures. To be able to grant an accredited degree and become eligible for federal funds, including federal loans to students, an institution is judged by a group of other institutions through a private association called an accrediting association. There are five of these private associations, regionally located, that provide the basic accreditation structure for higher education (non-profit) in the United States today. They set the guidelines for the process, appoint the visiting teams of individuals from accredited institutions who will make the report, and make basic judgments about the quality of education based on a self-study by the institution and a report by the visiting team. The final report itself is usually not made public, and the regional accrediting boards make their decisions in closed session.

Most universities, especially the larger ones, perform the ten-year review process in a rather perfunctory fashion, except that it does give political factions within a university the opportunity to lobby with the visiting team

to establish their priorities in the report. But smaller institutions are, in general, more vulnerable, and therefore must be more responsive to the desires of accreditation teams and the associations. However, if one is *imitating* the basic structures of courses, credits, Ph.D. faculty, and physical buildings filled with fulltime students, even if this is being done fifth rate, there is little difficulty with the accrediting associations. There is more difficulty, however, if one is attempting any substantial variation on any one of the major structures.

In recent years the more established and largely quantitative criteria for judgments of quality within the guidelines of the accrediting association, such as the ratio of Ph.D.'s to students in a given faculty, the square feet of space per pupil, the number of books in an institutional library, have given way to a softer and more qualitative approach. However, the net result has been to place the criteria for judgment in the hands of whomever is appointed to the visiting teams. They inevitably bring their own assumptions and the experiences of their own institutions, against which they compare and make judgments. The result is a multiplying of implicit and explicit assumptions about what a "good" institution looks like, and what quality higher education is. Time and again one finds examples of institutions making slight variations, and being forced to justify this diversity in terms of equivalents for traditional structures and functions. The real danger of the whole accreditation process in this view, then, is not the occasional horror story of a highly politicized visiting team, or the entrance of some personal likes or dislikes into the judgments for or against programs or institutions. It is the much more substantial, but subtle and all pervasive process of homogenization that results from a private association policing itself with its own members, using a set of implicit and explicit, but unquestioned, assumptions about the quality of higher education and the set way of structuring it whatever the results.

It is difficult to know what all this information collected by the accreditation process means, what judgments are made on it, or what influence is being exerted across the country in countless institutions, with their own resource allocation and future plans, by the recommendations of the associations. However, for the most part, institutions take very seriously the recommendations of the accrediting associations, and in general try hard to stay out of any serious conflict with the associations.

In addition, it should not be forgotten that the accreditation process

determines where federal monies can and cannot be spent on legitimate academic exercises. Judgments are clearly made based on some criteria about quality education. When the growing weight of federal dollars and eligibility for countless educational programs is tied to a process designed in other circumstances for other objectives, the questions increase. Is an institution really free to opt out of this private process? And if not, what are the prices paid by an individual institution, or higher education in general, or the public, by fitting all educational endeavors into these unexpressed but nonetheless rigid accreditation guidelines?

Since it became very clear in the 1950's that quantitative measures had questionable relationship to the quality of education, most of these were dropped. We now have a state of the art in which it would be extremely difficult to defend any kind of criteria that would in all cases automatically lead to quality education. To do so, one would have to argue that they must be present in order for an institution to be accredited, and, as we have noted, the associations refuse to do that. But at the same time we have a private and rather elaborate structure, which continually and quietly, but with great resources at its command, reinforces a single model of higher education for replication throughout the country. There is a great deal of pressure not to deviate from the commonly accepted structures, and anything that is a variation is viewed with strong suspicion. In their subtle, rather secretive, certainly unanswerable way, accrediting associations exist as a large and very important superordinate structure in higher education today.

Testing

The process of setting criteria for entrance into accredited colleges and universities is in many ways similar to that of institutional admission into the ranks of accredited bodies. Close to two million high school juniors and seniors take a national test, the results of which are used by the vast majority of institutions across the country, usually within some quantitative formula, to decide on admission and rejection of students to their institutions. The testing is designed by a private and untaxed corporation in Princeton, New Jersey, called the Educational Testing Service (ETS).[5] In addition to control-

[5]For an interesting "expose" of ETS, see Steven Brill's article "The Secrecy Behind the College Boards," October 7, 1974, *New York* magazine, pp. 67-83.

ling through a private association a very important tool in admission to colleges and universities, the ETS also administers tests to 74,000 students for admission into business schools and to 120,000 others yearly who wish to be lawyers. Others who wish to qualify for advanced placement, or who enter competition for the National Merit Scholarship, also compete on tests designed by ETS. Ignoring for a minute the near monopoly that ETS holds in an extremely profitable economic area, and its non-profit status with respect to this, the question that arises in relationship to the concerns of this essay is one of control over and above any individual institution in its own selection of students. Again, a private, non-regulated agency is providing most of the classification criteria for candidates for four-year colleges and graduate and professional schools. Both the accreditation process itself and the distribution of money are linked intimately by the fact that students who need financial aid to pay for post-secondary schooling, law school, graduate school or business school must have their "need" for the money verified by ETS before the schools can release the money to them. Often the money is federal or state dollars—public money.

In summary then, a private association, ETS, establishes the criteria, in terms of a series of national tests, for selection into or out of colleges and universities in the United States, and at the same time verifies the financial "needs" of students who wish to attend those institutions. This links to the federal guidelines, which require spending those monies only in accredited institutions. And the accreditation is controlled and administered by a private and secretive set of associations unanswerable to the public for judgments of quality of academic programs or responsiveness to client-student needs.

A perfect example of the conflict between public concerns and private associations in the area of testing has to do with racial and cultural bias. A white male from a high income family in the Northeast United States does best on ETS testing. There is a significant difference between the median scores of whites and non-whites on the majority of ETS testing. A private association has no direct responsibility to search its own values to determine whether the tests really are a mirror of the majority class skills and thought patterns, or whether the young of distinct cultural and racial background are being discriminated against. These are naturally *public* questions when one is making decisions about entrance into, or exclusion out of, higher education and the professions. And given the present set of private structures, they would appear to be extremely important questions.

In summary, then, the system of higher education gatekeeps with a series of tests whose development and control lies with ETS. Accredited institutions often award much of their financial help based not only on financial need but on some judgment of "merit" based upon the results of ETS testing. So the accrediting process, the availability of funds, the awarding of those funds based on testing, all are linked to private associations, unanswerable to the public, and also unanswerable to the individual institutions or students in any direct, meaningful way.

Certification

The final link in the chain is the process of certification of specific skills and licensing of certain occupations, and the ties between this very complicated process and accredited institutions. Over seven million people in the United States work in occupations that are licensed in one or more jurisdictions, and the list continues to grow. In the majority of cases, responsibility for licensing lies with the departments of state government, yet no two states license all the same professions. Licensing is obviously an attempt to regulate good practice and protect the public. But it also brings with it prestige, control of a profession, and greater job security. It is very often then, the occupational groups, and not the public, that demand government licensing, for it is ultimately a powerful exclusionary device.

The central criteria for most licensing is the proper educational background. In a credential-minded society this represents a single model approach to practically every occupation/profession. As the Newman Report to the Secretary of HEW in 1971 pointed out about the licensing procedure: "Rather than provide alternative routes, nearly all of these agencies reinforce the monopoly which educational institutions have over the awarding of credentials" (p. 40). What has happened over the years is a growing emphasis on formal educational experience as the requisite for licensing–based probably on an imitation of the elite professions. This schooling requirement then becomes tied, of course, to accredited programs of training. This in turn leads back to the process of the private interest groups accrediting, and therefore policing, themselves. For the individual state licenses only graduates of accredited programs, and this becomes the gatekeeping mechanism of the occupation/profession. The monopoly over entrance into specific occupations is delegated *de facto* by the state to private interest groups accrediting themselves. In addition, the relatedness of additional accredited schooling to skills needed in the profession are lost from public scrutiny.

The basic evaluation problem being dealt with here is the classic one: Do you evaluate behavioral skills related to the practice of the occupation/profession, or do you evaluate a program or process of formal education and then give automatic approval to any graduate? In the case of the law profession, for instance, although law boards are given as an entrance exam into the local professional ranks, the boards themselves vary greatly in their degree of difficulty. Therefore, reciprocity between and among states is chaotic. And even in this case of an exam in some cognitive skill areas, one must have finished a full course in the formal study of law at an accredited school beforehand.

Somewhere in the tensions between a meritocracy and a democracy, where there are some vague concerns about justice as fairness, a language of being "certified" and "accredited" appears to solve some of the problem. The line between the haves and the have nots is often maintained by simply moving the set of required certifications higher still. Licensing seems to have emerged as a prime tool in this process—not a protective one for the public, but a protective one for the profession/occupation.

A review of this entire set of processes and their interdependence leads one to an extraordinary conclusion. Surely there is a deep public concern about what educational institutions are declared "accredited," how public funds are spent in what places for what kind of study, how students are selected for colleges and universities, and how individuals are licensed to practice specific skills in this society. The entire question of what is legitimate to learn, how it can be learned, and perhaps most importantly, who has access to it, is contained within these interlocking structures. But there appears to be a strong case that these structures are not only not responsive to public needs or control, but that they form a tight circle of accreditation, testing, distribution of resources, and licensing, protected by autonomous, secretive, and ultimately unanswerable, private interest groups. The result is a closed model, unable to question its assumptions, methods, or direction. It represents a set of structures "outside" of any individual institution, yet dictating what occurs within. It is a set of controlling structures that is totally uncontrolled itself.

II. THE ACADEMIC PROFESSION

The Context

The network of formal structures that has grown up around higher education often comes as a surprise, especially when their private nature and interdependent relationships are made explicit. But they are only part of the story influencing questions of quality, evaluation, and public knowledge. The other important factor is the academic profession itself, with its training, emerging identity, and built-in assumptions about success, excellence, and prestige. One of the basic questions that lies at the heart of the ACE rankings is precisely the question of *how the academic profession measures itself.* As we shall see later on, the methodology employed in the studies selects out a specific group of deans and faculty members to express their opinion on "quality" programs. Since little or no criteria are provided on which to base their judgments, the results give strong clues about the assumptions that underlie the profession, its judgments about the "ideal model" of a department, and the academic profession's value system with its attitudes toward knowledge, learning, and success.

The professional attitudes are born and exist in a true sense outside of any individual institutions, and yet have a great deal of influence over the activities and goals of groups of academic professionals at colleges and universities across the country. Paradoxically, these attitudes in many ways have more control the farther one moves away from the core institutions that possess a majority of faculty that embody them. For instance, a faculty member at Harvard with a research reputation—a record of publishing to his or her credit—is taken for granted. However, as one moves out into institutions where there are only three or four such faculty in a group of three hundred, the power of the research prestige and publishing reputation magnify rather than reduce themselves. The "professional" ideal in many senses exercises more control on attitudes and choices the farther one gets from those who practice it.

The number of college and university professors in the United States in 1970 was 300,000 plus. They were at the center of the academic community in a country strongly influenced by both its natural and behavioral scientists. Until quite recently, the academician was really part of a small group of

individuals closely affiliated with the special mission of a particular institution. Allegiance was to *this* group, and only secondarily to the academic discipline, and even less to an overriding identity of an independent professional such as a lawyer or doctor or biologist. But this is changing as the training for a Ph.D. becomes standardized; as national publications and meetings create common values and shared information; as the job market evolves into a national market. These changes and the special type of unionization makes for a group of academicians with allegiance to their profession rather than to an institution or to a given employer/client. Professional rules and standardized practices, with attention on access and promotion, often grow narrower in such moments. And these, in turn, bring a profession face to face with questions of how it evaluates "quality" in its ranks and in its institutions. And the deepest assumptions and values of a profession emerge in the process.

One can begin a look at the academic profession by pointing out that at the roots of a teacher-learner relationship there is a process of evaluation. However, the way in which this process occurs requires some reflection. For whatever historical or pedagogical reasons, the acceptable practice in education is a type of "grading" by the teacher of the learner, most often not in terms of personal goals or degrees of content mastery, but over against peers through some comparison with "the group" or a norm. Throughout this society, the process is continual and all-pervasive. The classical sociological term is "meritocratic"—an attempt to evaluate people according to competence and achievement, rather than some other criteria. Jencks and Reisman see the growing professionalism of the academic in the U.S. as clearly increasing this kind of meritocracy. But for them it occurs almost inevitably, as a result of life in "highly organized societies with a very specialized division of labor."[1] The specialization has surely taken place, and so also has the attempt to apply a process of evaluation throughout the system. In the selection of graduate and undergraduate students, as well as the selection of their own colleagues, "merit," however defined, has become the academic yardstick. Perhaps it is not surprising, then, to find ready acceptance of a "ranking" of graduate schools. In fact, there is probably no other process quite so natural to the academic profession.

---

[1] C. Jencks and D. Riesman, *The Academic Revolution* (Doubleday & Co., 1968), p. 12.

This readiness to establish a pecking order is heightened by the fully autonomous nature of each college or university in the United States in terms of resources. There is no central planning agency that establishes priorities and allocates money or talent accordingly. Each institution competes on a national scale for faculty, students, and money with no visible future planning or laws of equal distribution. Any dysfunctions between supply and demand, or need and services produced, are more or less handled in a wide-open free market. It was probably this waste in training time and resources—without the most primitive manpower studies or needs assessments—that motivated Alan Cartter's study. As good as the motivation was, and as correct as his manpower projections proved to be, the choice of a methodology ranking "the best" in order to thin out the rest, raises serious questions. And in this openly competitive market, the award of national prestige by the American Council on Education (ACE) to the top twenty increases the seriousness of the mistake. The importance of this context of open competition—and the unexpressed, but operative, narrow, and elitist rules by which the game is played—is what the ACE studies manifest in such high relief.

The "Ideal" Model

Although quite often unexpressed, the process of evaluation in higher education is taking place within a context, as we have seen, of assumptions and values that underpin the academic profession. These assumptions, in turn, canonize an ideal model of the American university, and influence the growth of the profession itself. Talcott Parsons and Gerald Platt's latest book, *The American University*, is the most complete example available of the "ideal" model of the university that exerts so much power. The "full" university for them combines four descending functions:

(1) pure research;
(2) training future researchers and teachers;
(3) instructing undergraduates (as citizens);
(4) training practitioners in professional schools.

What is central to this model institution of higher education is its emphasis on the core value of *cognitive rationality*. This way of knowing has clear primacy over all others, although in the training of undergraduates and professionals it admittedly becomes somewhat diluted. There is little if

any question throughout the study by Parsons and Platt about the dysfunctions or conflicts within this kind of ideal model of a university. The idea that personal and academic skills that contribute to good teaching might be very different from those that contribute to research is ignored. A simple conflict in time allotment between what an institution demands of the professional in terms of teaching, and what the profession itself demands in terms of research, is sloughed off. There is the additional question of whether those who are good practitioners within individual professions are also those who do good research. And finally there is the question of distortion both to the institution and to the individual learner when the single model of learning described by the authors as cognitive rationality is held as an absolute. What occurs, of course, in such a description of the ideal model is a resultant elitism in which only a few special institutions of higher education can truly be described as "full universities." There is a certain kind of faculty whose primary activity is research, and they are really deemed to be the ideal kind of academic professional. The fact that this describes fewer than 20 per cent of the professionals practicing in institutions of higher education appears to be irrelevant.

Once the academic world is defined as resting upon the value of cognitive rationality, Parsons and Platt can begin a rather unique discussion of influence and prestige of institutions that deal in "intelligence." There is a remarkable chapter in *The American University* comparing the amount or capitalization of intelligence within a society to a banking model. For Parsons and Platt, the university is the intelligence bank of a society, where "cognitive solvency is relevant not only to evaluation of performances of borrowers, but to collective depositors. The intelligence entrusted to the university by students and faculty cannot be squandered on intellectual misadventures that leave the institution bankrupt with respect to its *previous relative, intellectual position in the total academic system*" (italics added).[2]

The above discussion is one of the more incredible quantitative discussions of the amount of "inflation and deflation" of intelligence within a society. But what follows from this discussion is that the relevant medium

---

[2]Talcott Parsons and Gerald Platt, *The American University* (Harvard University Press, 1973), p. 308.

for a university is influence, not knowledge or service. The authors describe influence in a rather interesting way in terms of the university. For them, *"influence is persuasion from position of relative to relevant prestige."*[3] The university then becomes an influence bank as well as an intelligence bank. Prestige is the capital that one must collect in order to spend it as influence. The model then leads to exactly where an elitist model would lead to an institution operating within the economics of influence, where the single most important currency is prestige.

It is extremely important to understand the subtle but direct tie between this definition of the ideal university as a research institute, based on a single model of learning, and national rankings by reputation. The model builds upon elite faculty and elite functions that automatically lead to questions of influence and prestige rather than, for example, to questions of increase in knowledge, service to the client, and so forth. It in no way describes the reality of higher education as it now exists. Nor does it open up the discussion to a question of what might be possible. It defines what *should be* in terms of *its* definitions of learning, knowledge, and service. In terms of the model's net effect on both the sociology of the academic profession and the future of research in higher education and its own self-assessment, Donald Light of Princeton University observed the following in a recent article in the *Sociology of Education.*

> Although Parsons is not persuasive about the functional integration of research and teaching nor about the theory of the ideal-type university, his writings do serve another function. If we ask as sociologists of knowledge why such ideas are written and accepted as scientific truth, it appears that they legitimate the "supremacy" of research universities. *This "theory" is offered as a general model for all higher education in America today when clearly it is not.* Even though research universities provide important benefits to society as a whole, they have costs and dysfunctions as well. In the guise of science, the arguments about "functional integration" and "ideal-type" universities are actually apologia for the dominance of this one group of institu-

---

[3]Parsons and Platt, p. 308.

tions. In this role, the "theory" is quite functional, and it will continue to be used by those whose interests it serves (italics added).[4]

Effect on the Profession

What is the effect on a profession of creating and idealizing a set of goals that only a minimum of its members can and do achieve? Its *de facto* use, as Light suggests, can justify the dominance of a certain few universities. But the canonization of "cognitive rationality" actually begs the question of research—what kind and about what subjects. The model describes a certain kind of research professor, and in its elitist definition implies that it accounts for *all* faculty. What the long-term effects of such obvious denial of reality will have on the profession is a difficult question to answer, but some early warning signals can be spotted.

Parsons and Platt's own data point up the significant discrepancy between what the members of the academic profession find themselves doing and what they would like to be doing. For instance, in "Graduate Teaching and Research," the difference between actual and ideal uses of the professionals' time at all three quality levels of institutions described by Parsons and Platt was alarming.[5]

AVERAGE ACTUAL AND IDEAL DISTRIBUTION OF TIME FOR GRADUATE TEACHING AND RESEARCH

	High	Medium	Low
Actual	50%	45%	15%
Ideal	65%	58%	47%

---

[4] D. Light, "Introduction: The Structure of Academic Professions," *The Sociology of Education*, Volume XLVII (1974), p. 10.

[5] Quoted in Light's article, p. 8.

There is a tension between teaching undergraduates and research even, and especially, at the "low" quality institutions. The contradictions, of course, run even deeper. Ph.D. training is almost entirely research oriented, and yet most graduates end up teaching. The institutions that employ them want them to teach, yet tend to promote them only for publication and research, or imitations thereof. It is often remarked that the most dangerous award on campus is the "best teacher of the year." Nothing arouses suspicion among peers more quickly and profoundly in a university today. The dangerous information is that this person is doing a better job pleasing the clients than other professionals in the department or university, and—by implication—that he or she is doing less research.

In the ideal model, it is clear that both the structure and focus of the university derive from the academic professionals. The pursuits of the university originate with the interests of the academic discipline, not the questions of the broader society nor the demands of the students. In fact, the role of the student in the model is not client but initiate. It is not the "education of an individual" but the "making of a professional" that emerges as the goal. However, as this country enters a stage of growing consumer interest and power, all the professions—including the recently emerged academic one—will probably be forced to recognize the client with something other than disdain and contempt. Both the focus of the research and the pedagogical delivery will have to respond more than at present. In the ideal model, and to a great extent in present-day practice—however schizophrenic—there is ample evidence that *teaching is not a professional activity*. One does not become known nationally for it;[6] it does not play an important part in professional promotion; and success in it can actually become a liability.

As one might expect from this discussion, almost all study of the academic profession begins with the question of research. Therefore, we know more and more about what 20 per cent of the profession is, and by a reverse process at least what 80 per cent is not. A recent article by Fulton and Trow adds to the growing knowledge about who really does research, at what institutions, and at what cost to other activities. Their research began with

---

[6]One of the worst fatalities of shrinking foundation resources was the cancellation of the Harbison Awards for teaching by the Danforth Foundation in 1973.

the critical question of faculty interest addressed to a broad range of institutions. In response to a question of whether their *interests* (not time allotments) lay primarily in teaching or research, the responses were the following:[7]

Very heavily in research	4%	24%
Both, but leaning to research	20%	
Both, but leaning to teaching	34%	77%
Very heavily in teaching	43%	

(N = 60,028)

Less than one fourth (24 per cent), then, favored research over teaching; 43 per cent favored teaching heavily, and more than three fourths (77 per cent) preferred teaching to research. This appears to be the distribution that is more and more substantiated by the growing research. In terms of actual time spent, Fulton and Trow use a division of institutions into seven "quality" categories from research universities to junior colleges. In the first ("highest") quality category of institutions, 70 per cent of the faculty spend six hours or less in the classroom a week, and 46 per cent of them spend four hours or less a week. On the other hand, in the middle and "low" quality four-year colleges, from 66 per cent to 75 per cent of the faculty spend from ten to twenty hours per week in the classroom.[8] It becomes clear that there is a great "division of labor" within the American system of higher education. There are small sectors where research totally predominates, and a larger sector where teaching is clearly the central function. This division between the activities of one fourth of the profession and the other three fourths highlights the difficulty of talking about a single profession with common activities and goals. If one adds to this the differences between the distinct disciplines, one realizes the impossibility of speaking of a single academic profession.

Absolute numbers of publications are often used to measure research skills, but that, of course, says nothing about the quality of that research. It

---

[7]Oliver Fulton and Martin Trow, "Research Activity in American Higher Education," *Sociology of Education*, Volume 47 (1974), p. 33.

[8]Fulton and Trow, pp. 38-39.

also says nothing about the different rhythms of the academic disciplines themselves. Fulton and Trow observe, for example, that "we still find a 25 percentage point difference in the proportions who have published between biology and the humanities."[9] What possible meaning can it have to compare number of publications between the humanities and science, or compare the way in which each uses one's time?

The data begins to show a greatly differentiated system of higher education, further complicated by distinct disciplines that require totally different activities, settings, rhythms, and research. But in the face of this complexity, the elite assumptions dominate the political reality and educational policy of higher education. The all-pervasiveness of the "norms" and their effect are summed up by Fulton and Trow's research in the following way:

> It is important to remember that although all higher education institutions teach, while only some do research, yet *it is research that confers elite status* on an institution and a discipline . . . the research-doing faculty embody the norms and carry on the activities which determine the character of the institution, and are chiefly responsible for its public image. *This group of active researchers dominates the leading institutions; and in many ways the research-oriented institutions dominate those lower in the system.*[10]

The potential for clearly defined and distinct educational offerings is there. But the departmental straight-jacket, and the professional norms, however unrealizable, demand the emulation of all. Rather than build on the diversity, the present course of the profession is to evaluate and reward a narrow spectrum of interest and activity, and ignore the serious distortion this creates.

Since this is the reality upon which the ideal is imposed, one can understand better the impact of the ACE rankings by prestige. They purport to validate the model (i.e., the nexus between research and prestige) and

---

[9] Fulton and Trow, p. 44.

[10] Fulton and Trow, p. 47.

strengthen the (rather questionable) professional norms. Although the question of educational excellence is still begged, the all-important political and financial effects of the rankings would be strengthened if the academic apparati could produce other validation. Warren Hagstrom delivered a paper at the 1970 meeting of the American Association of Advancement of Science that was later published in the journal, *Sociology of Education*. It attempted a series of correlations with the Cartter 1966 ratings for 125 departments of mathematics, physics, chemistry, and biology. Interestingly, the research was supported by a grant from the National Science Foundation.

Back-Up Research on the Model

The research was based on a multiple linear regression, and although there are problems in attempting to establish linear relationships with the Cartter data, this study still represents the strongest back-up research of the ACE rankings.

The key variables related to the ACE rankings were found to be departmental size, research production and faculty backgrounds. Departmental size by itself accounts for almost one third of the variance in the prestige rankings. There is no surprise here, since size is really the necessary but surely not sufficient underpinning of any of the other critical variables. Although "completeness" of a department varies from discipline to discipline, the larger-sized departments create the orthodox or total approach and can hire critical masses in sub-disciplines who can then perform large research projects with added prestige. In addition, it is *not* the correlation that holds much glory for Cartter, Hagstrom, and the others. It is a little embarrassing to spend large sums of money on elaborate rankings claiming a national identification of top quality, and then to discover that a single regression analysis on size of department produces pretty much the same result. It may be the strongest correlation, but no one spends too much time on it.

The more important consideration is that of "research production," which is measured in various ways. Hagstrom combines research article production and average number of citations to published research into one factor that accounts for more than half of the variance in the ACE studies. But this correlation is strong only when focused on the larger departments. When it is broadened out, it loses quite a bit of its predictability. It accounts for only

35 per cent of the variance in ranking of his full sample of 188 departments, in contrast to 52 per cent in the restricted sample.[11] It accentuates again the fact that a researcher is caught in a circular type of argument, where he or she is trying to make judgments of quality about *all* institutions while defining factors that only describe a special segment. The result is exactly what has already been circumscribed. And even after the researcher identifies some factors—such as size and quantity of publication—as correlates, the question of nexus between these and prestige, let alone quality, is still open to doubt. Hagstrom puts it well:

> It is clear that *no claims about causation can be made* on the basis of these results. Departmental prestige, the perceived quality of a department's faculty, *might very well be an important cause with respect to any of the "independent" variables considered here*. . . . Thus if one inverts the regression procedures used above, one can show that the inclusion of departmental prestige in regression equations can add significantly to the amount of variance explained by other vairables for such dependent variables as mean citations to published works, mean number of awards received, and mean amount of service on government advisory committees (italics added).[12]

This dilemma is built into any ranking not based on clear criteria. Any later attempt to "validate" the ratings is not able to do much but create other rankings of most citations, largest size department, etc. One may or may not get excited, for example, about the correlation with size, but no causal link can be claimed.

However unproductive the process of validation of these rankings is, the attempts go on, and they go on with a great need to prove at least the correlations. When this process fails, it brings not a little embarrassment, which leads one to ask whether or not this is a research question or an exercise in self-justification of the values of the academic profession itself. An article

---

[11] Hagstrom, "Inputs, Outputs, and the Prestige of University Science Departments," *Sociology of Education*, Volume 44 (1971), p. 383.

[12] Hagstrom, p. 398.

published in 1974 in the journal, *Sociology of Education*, is a good example of the tone of the present questioning about the ACE rankings. Using a combination of the Kenniston ratings of 1959 and the two ACE ratings, the authors attempted to construct a set of "quality of department" predictors to account for research productivity in sociology from 1940 to 1970. The question was whether or not the department where a student received doctoral training in sociology affected his or her scholarly output. The thesis was that the higher ACE-ranked departments would turn out more productive researchers. However, the results were a bit disconcerting:

> The correlation between the four ratings of quality and the six indices of productivity indicate only weak relationships between quality of department of doctoral training and publication output. In no instance does the independent variable account for even one tenth of the variation in productivity.[13]

That finding in itself is interesting, and the authors feel they must conclude that for sociologists, the quality of department of doctoral training as presently measured exerts little independent impact on publication productivity. However, what is alarming is the fact that the researchers never questioned whether the rankings really gave them "quality" judgments. The belief in the ideal model is so strong that the authors' discussion revolves only around how their methodology of linear regression may be too simplistic. The tone is one of disappointment for not proving the correlations, and a re-examination of their own methodology. The only question about the ACE ratings has to do with reputational methodology. The authors' remark:

> Anyone familiar with the literature on community power is familiar with the weaknesses of the reputational approach. Of course, from one perspective the results can be seen as casting doubt upon the validity of the ratings. . . .[14]

It was a tenuous suggestion after the fact, but at least it was the questioning

---

[13]Frank Clemente and Richard Sturgis, "Quality of Department of Doctoral Training on Research Productivity," *Sociology of Education*, Volume 47 (1974), p. 295.

[14]Clemente and Sturgis, p. 295.

of an assumption of the ideal model that we have found less and less examined since 1970.[15]

The case that seems all-important to make with respect to the present state of the academic profession and the impact of the ACE rankings is the following: the general context of higher education is extremely competitive; within this competitive structure there is a pecking order held in place by an ideal model of the university; that model is elitist in its description of the goals and central function of institutions of higher education as pure research; it describes the ability and interests of only a minority of the profession; yet the prestige of the minority of professionals, and a minority of institutions, continue to dominate the goals and directions of higher education. Hence the opening of the question of quality to other questions of diverse goals, methods, and settings is continually blocked. The ACE rankings are the single most powerful attempt to validate openly the assumptions about quality in the ideal model. As admittedly vacuous as their criteria of "reputation" were, there is now a constant attempt to validate the ratings with correlations with published research, citations, and, of all things, size. The danger with these exercises is the constant focus on only 20 per cent of the profession and its activities. The distortion that results to the profession itself, and to the delivery of higher education to students and the larger society, cannot be underestimated. *Rather than close off the ideal norms of the profession and the question of excellence, with an absurd listing of the top twenty departments the need is to open the questions of quality education, distinguish different goals and criteria for higher education, and make room for diverse roles within the academic profession.*

---

[15] See Appendix B.

III. THE STUDY ITSELF:

ASSUMPTIONS AND RHETORIC

The Origins

Simply to discover a ranking of academic departments–given the confusion as to goals and means, and the general atmosphere of resistance to any public assessment in higher education, especially graduate–requires an explanation. To begin with, there is the general American mania for ranking everything, since there's no use playing the game unless someone's number one. As Dr. Murray Eden, professor of electrical engineering at Massachusetts Institute of Technology, observed:

> Americans rank order everything–movies, books, ballplayers, restaurants, pollution, crimes, drugs, businesses, presidents, pimps and pederasts. I will not give a psycho-social explanation of the phenomenon. I state it as a fact simply. I think I know who benefits from these lists. They are of great value in reifying our American mythology; they provide useful information for certain occupational groups, and they are great time-savers. For example, you are on the Admissions Committee for the Epistemology Department. There is one applicant from Siwash who ranks 17th out of 131 in his class; and then there is another from Eastern A & M who was in the "upper seventh" of his class. Look up your Cartter report, convolve (as we engineers put it so precisely) the rank orderings and our work is done. Who needs to read the folder?[1]

Actually his conjecture is not far from identifying the origins of the American Council on Education (ACE) rankings. In the preface to the Cartter

---

[1]Personal letter dated March, 1974. During 1973 and 1974, Professor William Arrowsmith and I were in correspondence with many professionals across the country concerning their opinions on the ACE rankings. I will refer to their written responses throughout the remainder of the essay as "Letters to Arrowsmith and Dolan, 1973," etc.

Report in 1966, both Logan Wilson (then president of ACE) and Allan Cartter give the impression that the report was a result of planning interests of ACE's Commission on Plans and Objectives. However, Cartter does mention that the funding for the project came from the National Science Foundation (NSF), the National Institute of Health (NIH) and the United States Office of Education. The source of the true initiative for the study is important, since it provides some additional indications as to why such a study was desired, why it was approached in a certain way, and why rankings, specifically, emerged.

The rankings seem to have had their beginning over a lunch in Washington, D.C. in 1963 with Allan Cartter and Richard Bolt, who was then assistant director of the National Science Foundation. Logan Wilson, president of ACE, was also present. In the course of the conversation, Bolt informed Cartter and Wilson that the NSF felt the need for better ways of judging graduate programs, particularly since the amount of money was growing and there was mounting pressure to identify additional programs. NSF's basic selection mechanism had been the peer-group judgment of proposals, but they wanted to make it more efficient, sophisticated, and presumably more extensive. Bolt mentioned that NSF was thinking of asking a private research group to do an evaluation and *ranking of science departments*. At this point, in Cartter's words:

> Logan Wilson and I hit the ceiling and alerted him (Bolt) to the fact that there had been a long-standing position of the Association of Graduate Deans that *nobody* ought to play around with evaluating the graduate programs who was not themselves responsible to the institutions themselves; either the AAU or some other group. The deans had always said that anyone who is going to accredit us or rate us has to be responsible to the institutions. We don't want any outside body doing this. At that point I think we convinced Mr. Bolt that if it were going to be done, it ought to be done by a group like the Council of Graduate Schools or the American Council on Education or some other group.[2]

---

[2]Interview between Cartter and the author at New York University (February, 1972).

The basic message was the constant one that prevails in regional accreditation and professional groups in American education: "The profession evaluates the profession." The same protest was, in effect, made forcibly again here. And the result was predictable. Since Cartter had at that time been recently named vice president of ACE, and also director of its Commission on Plans and Objectives, this type of study looked like a natural way of answering some questions about the post-war expansion of graduate education and the effect of federal funding, while at the same time satisfying the NSF needs for a ranking of science departments.

Taking the initiative, Cartter returned to Bolt at the NSF with a proposal and budget of $105,000. The National Science Foundation put up $75,000, since it was their idea, and they went looking for the rest of the money from more broadly-based groups, since Cartter had enlarged the proposed study to include *all the major disciplines*, including the social sciences and the humanities. Politically, there was a clear need to get some wider support. They convinced the National Institutes of Health to contribute $15,000, and the Office of Education came in with $25,000. The Office of Education money seemed just enough to show that it was not strictly a "scientific" interest group, but not enough to alarm the institutions with what might be called a "federal" ranking.

So with some initiative from the NSF, hoping to get a better guide for spending its money, the result was a proposed numerical ranking of all major departments under the sponsorship of the prestigious American Council on Education. There was some resistance around Washington at first, but finally—after a great deal of hesitancy—people like Gustave Arlt, then president of the Council of Graduate Schools, joined the Advisory Board and gave their seal of approval. So the ranking had been arranged quietly by two "planners" seeking a bureaucratic tool for quicker evaluation, although the eventual rhetoric of the studies mentions aiding decisions by graduate students and administrators. It was clearly born, however, from the protectionist philosophy of evaluation so zealously guarded by professional education, and in this case its Washington lobby. It took its exact shape because the NSF was determined to have a list by rank, at least of science programs, and because of Cartter's own assumptions about "quality" graduate (and undergraduate) education and the need to identify it. His studies on shifting manpower needs were well ahead of their time, but the elitist underpinnings of his assumptions about quality seem to have led him to questionable instruments on which to base future planning.

The Assumptions

As we noted in the first chapters, there is general disagreement about the components of quality higher education, and, given this confusion, an obvious hesitancy to attempt to evaluate it. The ACE rankings, then, take the only route open to them. They ask deliberately vague questions about "quality" to form the ranking, and then, after the fact, attempt to establish correlations between the opinion survey and specific elements, such as publication, size of the department, etc. But again, the facile transition from the undefined question—with all its implicit assumptions—to an exact numerical ranking and encumbent statistical correlations is important to recognize and name.

Logan Wilson, in his foreword to the first study in 1966, declared that Allan Cartter's work was intended to further an understanding "of what quality signifies." However, when Cartter discusses "quality" in his preface immediately following, he ends by saying: "In an operational sense, quality is someone's subjective assessment, for there is no way of objectively measuring what is in essence an attribute of value."[3] The conclusion then seems to be that "quality" is what you (or a group) think it is, depending on what aims and means people have formed value judgments about. The statement is clearly an admission of the lack of agreed-upon criteria, as we have said, but when Cartter circles back with correlations with size of library, number of faculty awards, salary, and size, there is an attempt to make "someone's subjective assessment" *the* objective judgment.

The Rhetoric

A glance at the report's rhetoric and its uses is interesting, beginning in the foreword by Logan Wilson. One of the basic questions we will deal with throughout this essay is whether "hearsay and conjecture" change their nature or importance when collected and quantified. Ever since authors began to write, they have used rhetorical skills to lead readers to desired conclusions. Quasi-scientific and quantitative techniques are the latest and

[3]Cartter, *An Assessment*, p. 4.

perhaps the most subtle in this tradition. "Subtle" because the impression is given that something imprecise is made precise by an "objective mechanism" that transforms information apart from the author's subjectivity or control. Watch this sentence by Dr. Wilson, speaking of the decision by ACE to proceed with the rankings; it contains a justification and the critical transition from hearsay to objective numerical ranking.

> There was never any question about the need for doing in a systematic and objective way what necessarily goes on continually in any event, though usually in a piecemeal and more impressionistic way.[4]

It is somehow in the adding of "impressions" together in a "systematic and objective way" that the writer conveys a new result: a great many "subjective" conjectures make an "objective" one.

Since the study's title was *An Assessment of Quality in Graduate Education*, the burden of showing some correlation between reputation and quality would appear to be crucial. We shall watch this claim more in and out of the study, from the tables that list the rankings to their analysis and then to the attempted correlations. But Logan Wilson, perhaps better than we can, shows the distance between the two concepts in his defense of the study:

> To any who would discount the value of such an approach to the assessment of quality in graduate education, it should be added that in the final analysis the national reputation of a department or an institution is nothing more than an aggregation of individual opinions.[5]

Since that sentence appears clear even to those who undertook the study, the key question would seem to remain: Does this kind of opinion survey about reputation have any clear and direct relationship to specific elements of academic quality?

---

[4] Cartter, p. vii.

[5] Cartter, pp. vii-viii.

Further, the question of educational quality in the ACE rankings suffers the same confusion that exists at the heart of the accreditation dilemma. There is a vacillation between judgments of "excellence" on the one hand and minimal standards on the other. Because of the strong elitist assumptions of the study, both the methodology and the analysis are directed toward establishing benchmarks of excellence, or the "ideal model" of graduate education. In the process, however, a single type of institution is defined that begs the question of diversity of mission, resources, or students in graduate education. The need to create an elite, construct a single model of excellence, make the judgment of who can enter the club, and, as an afterthought, clarify the question of "quality" that really underlies the entire process, is nicely put in the first paragraph of the preface by Logan Wilson.

> Excellence, by definition, is a state only the few rather than the many can attain. Striving for academic excellence, however, is a worthy ideal for colleges and universities, and it can be reasonably argued that every educational institution should meet minimum qualitative standards, and particularly if it offers graduate work. A present problem is the need for a better general understanding of what quality signifies. Alan Cartter's book *An Assessment of Quality in Graduate Education* is intended to further such an understanding.[6]

The second study by Roose and Andersen significantly uses the much simpler title, *A Rating of Graduate Programs,* without reference to "quality." In explaining this change they say:

> One word about the book's title. Since it is evident in this study—as in the first study—that the appraisal is of faculties and programs as reflected *by their reputations rather than as they partake of specific components of an amorphous attribute called "quality,"* we have resolved to use as a title a simple description of the books contents (italics added).[7]

---

[6] Cartter, p. vii.

[7] Roose-Andersen, *A Rating,* p. xi.

Here is the clear admission by the official follow-up study that the rankings may mirror some opinion, but that the reader can gather next to nothing about the nature of "quality" in graduate education from that process.

Cartter's own preface continues the same tone that Wilson had begun. First of all, there is the rhetorical disclaimer about the pretenses of a modest study at the same time that he is praising the scope and profundity of the study. In Cartter's words, his commission had decided not to tackle an "over-all" review of higher education, "following the patterns of several earlier national commissions," but they would "study in depth selected critical areas."[8] It is hard to imagine something more "over-all" than a national ranking of all graduate departments in major disciplines, or anything less "in-depth" than a reputational study.

Continuing in his preface, Cartter attempts to calm any fears that this type of ranking is anything new by referring to two previous studies by Raymond Hughes—one in 1924 and another, sponsored by the ACE, in 1934. He also leans heavily on a third evaluation made by Hayward Keniston in 1957. However, these studies were substantially different attempts in much different circumstances. Hughes' second report classified graduate departments into two categories: adequate and distinguished. He used "expert panels" of forty faculty to rank the total of forty institutions; none of which were excluded. Keniston's goal was to assess the relative position of the University of Pennsylvania with respect to twenty-five similar institutions. He made his ranking by consulting the department chairmen at those institutions, and he reported his findings in an appendix to a study of graduate education at the University of Pennsylvania.

In contrast, Cartter, working with about four thousand raters, attempted a national ranking of twenty-nine academic disciplines, and he did so as vice president of the American Council on Education. With the ACE sponsorship went immense political clout and national exposure. As we shall see in the chapter on "The Use and Abuse of Statistical Hearsay," it is currently being used by state chancellors of education to justify more expenditures, by legislators to reduce monies, by faculty groups to recruit, and in student

---

[8]Cartter, p. ix.

handbooks to counsel students. It was a substantially different kind of study because of its scope and sponsorship, and because of the nature of the numerical rankings themselves.

In addition, of course, since the study claimed to be complete, any department not appearing in the ACE categories was automatically judged "not adequate." As we have seen, Cartter, with NSF prodding, was concerned with the rapid expansion of graduate education and was trying to put into the hands of policy makers at every level, a guide to aid them in cutting back on Ph.D. production. He counsels:

> Just as each college or university faces the perennial problem of whether to strengthen the weak academic areas or to build on present strengths, so for the educational system as a whole decisions of a similar nature are constantly arising which must be settled at the state, regional, or national level.[9]

As correct as his forecast of over-supply was, the manner of deciding who was "outstanding" and who was "not adequate" comes down to a glib admission that "quality" is simply " what the experts ranked." Roose and Andersen in the second ACE study three years later were so uncomfortable with this claim that they evaded the quality question, at least in nomenclature, as much as possible.

Once the origin of the rankings becomes clear, several aspects of the final shape also are more easily understandable. Some of the more implicit assumptions about quality, depending to a great extent on the state of the academic profession, have also been singled out as present and operative in many of the judgments behind the rankings. Although it flows out of much of the general academic "professional" posturing, it appears to be especially prevalent in these studies and influences them in many ways. The concern here is not directly with the ethical/political dimension of elitism versus populism, but with a series of unquestioned assumptions about where good education takes place, and more importantly, who can and should have access to it. The eventual effect of such assumptions is to make any investigation

---

[9]Cartter, p. 8.

based upon them narrow to the point of being myopic.

Cartter, and the basic ACE position, appear locked into a single model of achieving "quality higher education." Diversity, then, can only be a question of better or worse in terms of the single set of criteria. Any question of real diversity for achieving quality education is incomprehensible in this type of closed elitism. It allows Cartter to unabashedly proclaim that the "poorest" graduates from the "selective" colleges are better educated than "the best students at some mediocre institutions."[10] Cartter clearly believes he knows what true "quality" in graduate education is, and therefore any flexibility or adaptability are automatically going to result in inferior institutions. The steps from these suppositions to his study and its purpose are outlined in this paragraph from his overview at the beginning of the study:

> Our present system works fairly well because most students and parents and prospective employers know that a bachelor's degree from Harvard, Stanford, Swarthmore, or Reed is ordinarily a better indication of ability and accomplishment than a bachelor's degree from Melrose A & M or Siwash College. Even if no formal studies were ever undertaken, there is always a grapevine at work to supply impressionistic evaluations. However, evaluation by rumor and word of mouth is far from satisfactory, particularly in advanced training for scholarship and the professions.[11]

The argument goes something like this:

1. There is a quality way of education.
2. We all know what that is.
3. But it's humiliating to rely on "rumor and word of mouth" to communicate this knowledge;
4. Especially in graduate and professional education;
5. Therefore, let's get these "impressionistic evaluations" tabulated, circulated, "objectified."
6. Therefore, the Cartter Report.

---

[10] Cartter, p. 3.

[11] Cartter, p. 3.

Even when moving to the seemingly more altruistic of his goals, that of providing information for prospective graduate students, Cartter is at a loss to cover his tracks. It's important, he says, to identify the lower ranking institutions too, for:

> Many prospective graduate students would not be suited to an education at Harvard, the Rockefeller Institute, or California Institute of Technology. Other institutions in view of their educational offerings, level, or work, and quality of students, would provide a happier and more productive experience.[12]

If this were a real concern, and really diverse means could be truly described, then the way to approach such a selection problem would seem to be a description of different "educational offerings, levels of work, and so on," and the worst way would be to rank the top twenty departments, and disregard half of the rest as inadequate and not worth mentioning. The latter method stresses the criteria of institutional size and resources; the former, if it were done, would be based on a realization of different kinds of students, institutions, and processes, with emphasis on the need for diversity and correct matching.

In the Roose-Andersen study, although the methodology is a replication of the Cartter study, the emphasis has shifted, as we have begun to note. There is much less talk of "quality" or concern with establishing an elite pecking order. By 1969 the emphasis at ACE had swung away from a focus on the identification of excellence, and, not surprisingly, is at the other pole of assessment—the establishment of bottom-line data for consumer choice. The authors, on reflecting on the objectives of their study, declare that their new focus is on the need to "protect the potential consumer of graduate education from inadequate programs." The authors immediately move to the obvious question of what information is really supplied to a consumer by a numerical ranking based on vague criteria:

> Once a faculty and program are acknowledged to be of a quality that merits student enrollment, many factors other than a

---

[12] Cartter, p. 3.

rank order or specific score should be considered by the prospective student.[13]

In terms of this objective, then—of drawing the base line of, for instance, sixty programs in psychology—a prospective student could not have chosen a more inappropriate instrument than Cartter's reputational ranking of the top twenty. Instead, the student would look for some development of the "many factors" that should really bear on student choice. Roose and Andersen themselves mention a rather appropriate beginning list, such as existence and quality of a particular specialization, breadth of program, availability of faculty, and so forth. But the authors could not repeat the same kind of ranking by reputation as Cartter used and pretend it would serve a different and legitimate need by simply saying that is what they wished it to achieve. The rankings do undoubtedly find their way in an all-pervasive manner into the student guides, but what has been disseminated in the name of consumer protection? An assessment that not only did not have student criteria built in, but also left students out as respondents in any way.

The other basic decisions that underlie the ACE studies and that are so much a part of higher education that they scarcely draw notice, are worth noting. Those are the choices to rank graduate *education* by academic *department*. These are two perfectly natural categories that are obviously in need of assessment, it is true. But the type of assessment and its context cannot be separated from these two choices. Again, the reader must recall that what results is a public list—published in rank order by the American Council on Education—with emphasis on the "top twenty." The public list appears in a context similar to a vacuum, which has no national evaluation, prestigious or not, based on clear criteria for judgment. The result is a powerful national dissemination of the American Council on Education's "objective" judgment of the "best universities." That is the reality—both politically and educationally.

In this process two things occur. First, the large, orthodox academic departments are rewarded, and bring prestige to their universities. If an academic department, for whatever reason, had decided to concentrate

---

[13] Roose-Andersen, p. 2.

strength in certain areas and not build in the orthodox fashion, it would suffer. If interdisciplinary work had begun, the departments involved would suffer on such a ranking; the interdisciplinary activity would not appear on any of the rankings even though it might be some of the most exciting graduate education occurring, and the university would realize it had been penalized for experimentation. No matter how strong and clear-minded a university is about its own set of priorities and missions, this type of pressure is bound to have enormous effect on planning, experimentation, and future direction of graduate education. Only the most naive could deny the powerful limits such a departmental ranking places on graduate education.

The impact on undergraduate education is an obvious result of its being totally ignored. The response that this is a graduate ranking only does not console the universities left off (if they have graduate programs), and it conveys to all universities that their reputations have little if anything to do with what is occurring at the undergraduate level. By the time of the second ACE study in 1969, Roose and Andersen were very much aware of this problem.

> If, for example, the response to these rankings is to shift resources and attention in excessive amounts to the development of the quality of graduate programs, undergraduate teaching and education could be the ultimate losers.[14]

But the *de facto* situation is that the power and prestige of the ACE rankings *do* have precisely the effect that Roose and Andersen fear. Undergraduate excellence receives no recognition, brings no prestige, and therefore is bound to lose resources. In addition, as we have noted, what is being rewarded in the ranking is probably damaging to good undergraduate education. The end result of the choice to rank graduate departments was to emphasize the full-scale research university and make it the ideal model of quality education.

It is difficult if not impossible to really get a glimpse of the prestige academic departments from the "outside." The two broad observations—about narrow orthodoxy and the unimportance of students—that are mirrored in the rankings can be made solely from the data of the approach of the study.

---

[14] Roose-Andersen.

its methodology, respondents, and assumptions. But the reader yearns for specific documentation of quality—of which there is little, not unexpectedly—with the conviction that a description of the real educational environment would be even a greater surprise. On occasion, however, a tip of the reality does surface, with enough information to confirm the worst suspicions. A recent example does say something to our argument and raise questions about the need for public disclosure around clear criteria for educational quality, such as contact with faculty, breadth of offerings and schools of thought, combination of theory and practice, and so forth.

The department of economics at Harvard University ranked number one in all categories in the ACE 1964 study and in most categories in the 1969 study. *Time* magazine published a story in February of 1975 describing the present environment of that department. One of its Nobel prizewinners was resigning, describing the Harvard economics department's curriculum as too narrow, and observing that the faculty had lost touch with the students. Many of the students found the focus was on theoretical mathematical model-building to the exclusion of all else. A visiting committee of trustees had filed a report within the last year that described the first year curriculum as "disgraceful," and the whole focus of the department as outmoded and in need of reform. The Harvard Trustees' own report found the senior faculty "inaccessible and unapproachable," and the student-faculty relationship filled with mistrust.[15] The trustees report was not distributed for a period of eight months, until the results were leaked to the student newspaper.

So the *number one* ranked economics department with sixty-plus faculty appeared to be educationally narrow and oppressive, inaccessible to students, and determined not to deal with any evaluative information about itself or its work. Perhaps, however, the ACE rankings do work to a degree, for the Harvard economics department slipped a rank to *number two* in the second category of "effectiveness of graduate programs," in the second study.

To examine how the ACE studies positioned themselves as they approached the question of evaluation of higher education is important to note.

---

[15] *Time*, February 17, 1975, pp. 77-78.

The motivation for the study does more to explain the "ranking" methodology than any later justification. The assumptions about excellence, quality, the "top twenty," are all present before going into the study itself and they affect its outcomes. The rhetorical accoutrements of statistical validation and other claims to "objective and scientific" outcomes based on subjective hearsay begin as early as the two prefaces. The confusion about the role of evaluation that exists in judgments of accreditation, eligibility, and certification, also exists in the two ACE studies, to further cloud the goals. The judgments to focus on graduate departments with this national ranking automatically affect the strength and importance of (1) undergraduate education, (2) interdisciplinary work and (3) experimentation in graduate education in general. The use of resources in these three areas penalize universities in the national ranking based on prestige. Finally, when one catches a glimpse of the reality behind a number one ranked department, as in the case of the economics department at Harvard, one has to question the validity of such a process as the ACE rankings and the significance of their results.

IV. THE STUDY CONTINUED:

METHODOLOGY AND CONTENT

The approach to any question is conditioned by the assumptions brought to it.

The assumptions–that become clearer as the rhetoric of the ACE reports is examined, as well as the origins of the reports–were bound to have their influence on the methodology, and hence the results, of the studies themselves. How the *questions* were phrased had a great deal to do with the assumptions about "quality" and the implicit criteria contained therein. The choice of *correlatives* to strengthen the rankings and to give increased "objectivity" followed these assumptions. The choice of the *department* as the unit to be measured depended, as we saw, on the origins of the studies and the desires of the sponsoring agencies. The focus on prestige, and prejudices about "elite" institutions, led quite naturally to the choice of *reputational studies* and the manner of choosing the universe of correspondents. It now becomes abundantly clear how the choice of methodology began to influence the content of the studies and their results. This chapter deals with the links in that process.

The Questions

Rather than an elaborate questionnaire, Allan Cartter asked only two simple questions in the first study, leaving open to the correspondents whatever meaning they chose to give to the words "quality" and "effectiveness of doctoral program." However, some clues are given in this sparse approach and they become quite important. First the two questions:

> (A) Which of the terms below best describes your judgment of the quality of the graduate faculty in your field at each institution listed? Consider only the scholarly competence and achievements of the present faculty. Limit the number of "distinguished" to no more than five.
>
> (Distinguished, strong, good, adequate, marginal, not sufficient to provide acceptable doctoral training, insufficient information.)

(B) How would you rate the institutions below if you were selecting a graduate school to work for a doctorate in your field today? Take into account the accessibility of faculty and their scholarly competence, curricula, educational and research facilities, the quality of graduate students and other factors which contribute to the effectiveness of the doctoral program.

(Extremely attractive, attractive, acceptable, not attractive, insufficient information.)[1]

The first question calls for a judgment of faculty with strong emphasis ("consider only") on *scholarly competence.* Although we will look much more closely later at who is doing the rating, and what that might tell us about what they are rating, it seems hard to know at this point what could be rated by the first question, other than publications and perhaps the overall reputation of the institution in which the given department was located—in a kind of halo-effect. Cartter immediately limited the categories by calling for only five "distinguished" rankings, which, of course, was consistent with his purposes and assumptions, as we have seen. By doing so, he also communicated to the rater what kind of universe and distribution was really being sought. As Dr. James I. Armstrong, president of Middlebury College in Middlebury, Vermont, pointed out with regard to "limiting the universe":

> I note in the Roose-Andersen questionnaire that the respondent must limit the number of distinguished ratings to no more than five. I simply don't understand what that means except that a prior "practical" judgment has been made that surely there cannot be more than five "distinguished" graduate departments in classics, etc. This may be true, but it appears to me to be a condition imposed by the method rather than a condition designed to seek the truth whatever the truth may be in this case.[2]

The second question attempted to turn the focus from a judgment of faculty scholarship to the "effectiveness of the doctoral program." Cartter

---

[1] Cartter, *An Assessment,* p. 12.

[2] Letters to Arrowsmith and Dolan, 1973.

has stated that he included the second question "to avoid the charge that the present study was just a popularity poll."[3] It is extremely difficult to see how this second question might avoid that result, but he mentioned in the same explanation that he was attempting to separate "their assessment of the quality of the scholarly contributions of the faculty from their assessment of the educational program provided doctoral students." However, the correlation between the two questions is .99, and one could reason from such a correlation that there was only one true variable. Even the very well-informed academic—who would know who was doing the top research and where—would have very little if any recent knowledge about the calibre of doctoral programs, accessibility of faculty, and teaching skills. Regarding this question, Dr. Frederick Crews, professor of English at the University of California, Berkeley, remarked:

> I do believe there is something farcical about ranking colleges as if they were the Dolphins and Redskins down through the Oilers. My own department comes out looking fine, and three years from now it will undoubtedly plummet in the ratings, thanks to the retirement of virtually all our superstars; but our effectiveness as a department won't necessarily have changed at all. It is notorious, of course, that great men get a bit tired and disillusioned toward the end of their careers, and by then they are so ensheltered in privileges that the students have a hard time even meeting them. Probably most of the important teaching here is now being done by assistant and associate professors younger than myself at the age of 40. The rankings are a high-class form of gossip and nothing else; and the people who form the backbone of our department are just beginning to be gossiped about.[4]

If the answers to the first question represent only a modicum of informed opinion, the responses to the second must surely represent guesses made in ignorance.

---

[3]Cartter, p. 9.

[4]Letters to Arrowsmith and Dolan, 1973.

The Correlations

In order to strengthen the believability of the rankings, Cartter chose four disciplines in the first study on which to run a fairly complex set of correlations. One of the things he discovered, not to anyone's surprise, is that "there is a clear correlation between reputation of a department and the scholarly production of its members."[5] In economics, for instance, he found the following correlation between a "publications index" and the rankings.

> The ten most productive departments (nine of which are also among the first ten in the subject measure of faculty quality) accounted for 56 per cent of all publications reviewed for this four-year period; the 25 most productive departments accounted for nearly 90 per cent.[6]

Obviously, since the judgment of general scholarly competence is the underpinning of the ranking, the "publication in journals production of specific departments provides high correlation. The journals are the major outlet for academic publishing and are viewed as the most important single source of information about movement within an academic discipline. Therefore this correlation might be taken as a confirmatory piece of evidence that the rankings do indeed mirror rather exactly that amorphous attribute called "quality higher education." This could be true despite the very great variance in quality of journals within a field, and the frequent judgments about the poor quality of most of what is published. Still, the central question in this essay is whether we are continually dealing with a closed circle of professional judgments and assumptions, of which the rankings represent the most crass example. Do the journals really represent an outside criteria agianst which the ranking can be validated, or is the journal publication process itself an exact mirror of the politics of academic prestige trumpeted by the ACE rankings.

It is interesting to examine some of the major academic journals, their

---

[5]Cartter, p. 81.

[6]Cartter, p. 81.

"home" universities and the make-up of their board of directors, for clues to the process of academic publication. For instance, the *Journal of Political Economy*, published at the University of Chicago, shows 13 per cent University of Chicago authors from 1950-1969; and 1 per cent from Harvard University. However, for Harvard University's comparable *Quarterly Journal of Economics* the figures interestingly enough are reversed: 1.6 per cent Chicago authors; 13.4 per cent Harvard. A look at sociology begins to give the same impression. D.H. Shranklin, in an article in the *American Sociologist* on "Prestige and the Sociology Establishment" in 1970, pointed out that the *American Journal of Sociology* with University of Chicago sponsorship has been dominated by Chicago faculty and their recent graduates. *Social Forces*, with a North Carolina editor, is heavily weighted with North Carolina authors.[7] The major association journal, *American Sociological Review*, is so dominated by the pecking order that 82 per cent of the editorial board have been staff or graduates of the "top" five departments.[8]

The fairness of the journal system is the question. The policies of selection of articles in and out begin to sound faintly familiar. The prestigious journals are often controlled by an individual school; boards of editors are made up of a majority from less than 2 per cent of the degree-granting departments. Yet the publication in these journals is clearly the central way for the individual, and thus the department, to arrive. The prestigious departments appear to control the main avenue to prestige! If this were true, or if it were only half true, what would it mean to the circularity of the process, the conscious and/or unconscious exclusion and closedness of a system, the operation of a politics of the elite—even in the reputedly open forum of academic publishing? The correlation between the publication "index" and the rankings then may be nothing more than mirror images of each other, and hardly validation for each other.

The power of the prestige departments in turn depends a great deal upon the simple volume of the academic unit—its size and economic resources.

---

[7] D.H. Shranklin, "Prestige and the Sociology Establishment," *American Sociologist*, 5 (1970), pp. 154-56.

[8] D. Crane, "The Gatekeepers of Trainers: Some Factors Affecting the Selection of Articles for Scientific Journals," *American Sociologist*, 2 (1967), pp. 195-201.

Dr. John G. Kemeny, president of Dartmouth College and an eminent mathematician, discussed the size distortion in the following way:

> First of all, let us consider what the rankings measure. They in no sense measure any objective quality of the graduate departments in mathematics, for instance, but rather how these departments are perceived by the leading professionals in the country. A quick look at the ranking of the top twenty institutions in any of the surveys will show that they tend to be almost identical with the list of the twenty largest institutions, either in the terms of the size of the research staff in mathematics or in terms of Ph.D. output. In other words, they represent "the most *physical* institutions." I am quite sure that it is not a coincidence that most of the respondents to this survey will have received their Ph.D.'s from one of these institutions.[9]

Regarding the correlation of ranking with size of academic departments, Charles Elton has written two articles analyzing the Roose-Andersen results in 1970 in both physics and psychology. Using *A Guide to Graduate Study*, edited by J. Graman in 1965, and, ironically, published by the ACE, he has taken several "size" variables for each graduate program in physics and psychology, and "predicted" almost the exact distribution into categories that corresponds to the Roose-Andersen distribution into "extremely attractive," "attractive," and so forth. He used published data of (1) the number of areas of specialization within a department; (2) number of faculty; (3) number of Ph.D.'s awarded between 1960-64; (4) number of full-time students; (5) number of first-year students; (6) ratio of part-time to full-time students. The third variable (number of Ph.D.'s) best differentiates "the departmental categories in the ratings of physics departments," Elton found.[10] In psychology, using just the second variable (number of faculty), Elton was able to differentiate all the "extremely attractive" departments from all other categories, and the addition of variable six (ratio of part-time to full-time students) differentiated all the "attractive" and "acceptable plus" categories from the

---

[9]Letter to Arrowsmith and Dolan, 1973.

[10]Charles Elton and S.A. Rodgers, "Physics Department Ratings: Another Evaluation," *Science*, No. 174 (1971), p. 568.

rest of the universe of "less than acceptable" departments.[11]

A few of the results of the Elton articles give an indication that *"size" may be the central variable* consciously or unconsciously used by the raters. The confusion of "quality" with "quantity" seems apparent, and the clear danger is that there is surely at present a relationship between size of department and financial support. With the publication of the rankings of the ACE reports, the financial support could increase or decrease according to ranking. Thus, a dynamic would be assured in which the larger get stronger and the smaller get weaker.

In addition, there is the implicit message to those "programs moving toward quality graduate education" that Roose and Andersen seek to identify, that the bigger you are, the better you are. The institutions and departments are getting the message loud and clear. All this in the face of strong evidence that real "quality" education takes place in the sub-fields of a discipline in a department—that spreading itself into every branch instead of concentrating in several quality areas of a discipline is the worst possible priority to hold out to a growing department, especially in a receding market. Both these implicit messages are contradictions of Cartter's central concern for limiting the growth of graduate education, and they are a forceful eliminating factor for the smaller departments.

The Department

This discussion leads to a much more fundamental question that is

---

[11] Charles Elton and H.A. Rose, "What are the Ratings Rating?" *American Psychologist* (March, 1972), p. 199.

[12] It is interesting in view of these assumptions to see what Cartter says of ideal growth patterns for a department: "The concept of optimal size, however, should be applied to departments rather than to the university as a whole, since the major economies lie in the areas of distinguished faculty and library and laboratory resources. An outstanding physics department of relatively small size may require the presence of six or eight distinguished senior physicists, an excellent research library, and expensive equipment. An outstanding department with five times as many graduate students need not multiply these resources by five, however. . . ." (Cartter, p. 114).

begged in the ACE rankings. Is the "department" the unit that *ought* to be looked at when asking the "quality" question? It is clearly the major political /structural unit within the university, but there is almost universal agreement–even from those who do not mind the rankings–that quality of graduate education can only be discussed in terms of the cognitive subdivisions of the academic discipline or department. Dr. Murray Eden of MIT puts it this way:

> These reports are subversive of scholarship because they simply espouse the falsehood that the department is the intellectual unit that distinguishes one school from another. This is not true even for departments of physics or chemistry, disciplines in which we now have a tradition of large group research, because the domain of study is a great deal narrower than "chemistry"; for example, synthesis of organo-metallic polymers. The departmental lump needs to be looked at with a finer grain.[13]

Of course, a study that had aimed at qualitative description of academic departments could have dealt with this diversity of quality within a given department, but a quantitative methodology of the kind applied here cannot make this basic distinction.

However, this question of varying degrees of quality within the subgroupings of an academic department only opens the door to a more substantial set of problems. It is another dimension of the diversity question again, but with more bite in it. The methodology of the rankings presupposes that every department has the same focus and identical objectives, and therefore that the same vague criteria of "quality" can be applied. But how far is this from a diversity of quality departments–each pushing to achieve its own particular emphasis and goals? It is an allowance for this kind of diversity that is totally lacking in the ACE rankings, and the basic question is: What damage is continually being done by the string of rankings begun by Cartter and the ACE? Leo Treitler, a dean at Brandeis University and a member of the music faculty, pointed–as many others have–to this glaring weakness in the ACE rankings?

---

[13] Letter to Arrowsmith and Dolan, 1973.

> To begin with, the survey shows ignorance, and it does so at two levels. At the crudest level, it throws together departments that do not really have the same objectives at all, *although they are all called departments of music*. For example, Brandeis, [with] a very small department offering degrees in musicology and composition, is placed between Rochester and Southern California, two huge departments, both of them essentially conservatories for performers, with some teaching of composition and some service courses in history. Rochester's distinction and attraction is based on its training of performers; that of Brandeis is based on the scholarship and composition of its faculty and of the students that finish here. *The juxtaposition, therefore, has simply no meaning.* This absurdity is repeated throughout the list, and I don't think that my claims about it will be very controversial.[14]

If this type of numerical ranking is built around the unit of the department, and the researcher is unwilling or unable to define the characteristics to be judged, the effects become increasingly obvious. The sub-divisions of the department—where the true academic activity is or is not taking place—are passed over. Departments with substantially different emphases and goals are lumped into leveling vagueness. And if this is true of an individual department that is building a diverse concentration *within a discipline*, what of the more exciting *interdisciplinary* attempts that are severely penalized by the unquestioned assumptions in the rankings? Even the more traditional interdisciplinary departures—such as American Studies—don't appear in the ACE categories as legitimate. What is the effect, then, on the important "quality" concentrations of a new interdisciplinary type. Dr. Crews of Berkeley gives an example:

> I can speak concretely about the silliness of the ratings in one place. The English Department at Buffalo isn't even mentioned as "adequate plus" in any of the rankings, yet for students who want to specialize in my own hobby-horse, literature and psychology, Buffalo is virtually the *only* place to go. No

---

[14] Letters to Arrowsmith and Dolan, 1973.

other university has a fully integrated doctoral program in this field. Furthermore, the relationship between faculty and graduate students there is one of unparalleled intimacy and collaboration, whereas some of the high ranked departments are reputed to have Kafkaesque atmospheres.[15]

In deliberately developing an interdisciplinary program such as this, an institution like Buffalo runs the risk that both cooperating departments will be penalized, since the program represents substantial departures from orthodoxy in each of the disciplines.

There is a continual suspicion that even *within* the departmental rankings themselves there is an exclusionary process in the name of orthodoxy. That is, the reigning orthodox view—especially within the social sciences, but elsewhere as well—is being narrowly voted for in the rankings, to the exclusion of any department that has dared to venture away from this central disciplinary position, or at least attempts to contain some diversity or views on either the methodology or the future direction of the discipline. These suspicions are mirrored in the comments of Seymour Sarason, professor of psychology at Yale University:

> In the case of psychology the ratings clearly reflect a continuing attempt to ape the methods and presumed values of the natural sciences. That is to say, the so-called best departments are those perceived as being quantitative and experimental. Put in another way, the best departments are those in which quantitative methodologies are revered over scope and breadth of substance. Skinner's impact, mathematical models as a speciality, the view that one can only know or prove something by experiment—these are some of the characteristics of American psychology which the so-called best departments have in abundance.[16]

The general effect is obviously to further freeze the "orthodox" view within the specific discipline, and drive any variation to the periphery. The

---

[15] Letters to Arrowsmith and Dolan, 1973.

[16] Letters to Arrowsmith and Dolan, 1972.

serious question of the source of new directions and discoveries within academic disciplines then becomes extremely important. The constant pressure is to conform, to replicate, to ape the leaders, or run the danger of ostracism. The deadly result is department after department with identical distribution into speciality areas, with identical requirements and course titles, and with identical thoughts and research being produced. Dr. Jerry Glen, professor of German at the University of Cincinnati, had this to say about the rankings:

> Most German Departments, judging by their requirements, seem to feel that a knowledge of old high German is more important for a student specializing in 19th century literature than the knowledge of Shakespeare and Homer. (And in some cases, that a knowledge of Thomas Mann is more important for a student of old high German than is a good knowledge of Latin or French.) Unranked departments which might be interested in questioning these priorities don't have a chance. The ACE rankings unquestionably reinforce the status quo.[17]

The question here has been to examine the origins of these rankings and to make the assumptions behind them explicit. One can then understand why numerical rankings by departments appear with national ACE sponsorship. Once this is clear, the question can then turn to what was truly being ranked–size, publications, and so forth. Finally, a few of the effects of the rankings were examined with respect to diversity within academic disciplines. The constant question is: Who is tabulating these opinions, based on what assumptions, and what effect does the process have on graduate education?

The 'Public Need'

Hidden beneath the total context of the ACE studies is a curious contradiction that seems built into the history of the professions in the United States. The "public's" need in health care or legal services or education is met, not by the federal, state, or local agencies that spend the public's money on programs and products in these areas, but by the desires and judgments of

---

[17] Letters to Arrowsmith and Dolan, 1973.

the members of the respective professions. Quite often, and understandably, so, this process leads to major policy decisions that serve not the public's need, but the particular profession's priorities. The system of evaluation in higher education by private associations of the institutions themselves, which we saw Cartter defend so strongly while arguing for the ACE study, is a case in point. Logan Wilson, in the preface to Cartter's study, mentions three broad purposes of the research. The third one is "*to examine critically* the available techniques of evaluation."[18] One is therefore very surprised to find the same kind of evaluation taking place in the Cartter report itself. It appears that we are presented with the opinions of a highly-selected "club" of professionals whose vague judgments *about themselves* is put forth as the guide for public authorities and policy-makers to evaluate present programs and formulate plans for the future of higher education. It is somehow like asking the top executives at Ford and General Motors to rate their products against other means of transportation to decide how the public funds should be allocated in the area of transportation. The cry that "none can judge the professional but the professional," somehow results in a belief that we then possess the disinterested, objective truth from "knowledgeable respondents."[19]

Cartter does everything in his power to perpetuate this view. In his conclusion, emphasizing how helpful the ranking should be for administrators, he gives his view of what kind of judgment is given in the study:

> The present study does not pretend to reflect all aspects of eminence or its absence, but it provides an approximate view by one hundred or more *dispassionate judges* who are not on the university's own payroll.[20]

---

[18]Cartter, p. vii.

[19]Forrest Tyler makes this point in his article "Knowledgeable Respondents," *American Psychologist* (March, 1972), p. 193. He concludes: "So again we face the paradox that was apparently overlooked in designing the ACE study. *How can we evaluate ourselves as disinterested experts functioning in the interests of the public when we are prisoners of our own perceptions and seemingly unaware of that fact.*"

[20]Cartter, p. 118.

They are, however, clearly on the payroll of a certain interest-group within a given profession, and their judgment could hardly be termed "dispassionate" with regard to graduate education. The nature of the selection process itself, and reputational studies in particular, shed some light on the make-up of the "hundred or more dispassionate judges."

Reputational Studies

Both the ACE reports are what would have to be called "statistically complete." As we have seen, they each spend a great deal of time and effort with correlation after correlation between their rankings and such "objective" measures as the size of library holdings, wages, publications, and so forth. The correlations may be used less to prove that the ranking is correct, than to show what the reputational judgments were based on. We have also seen that with some rudimentary statistics about the size of faculty and student body and faculty salary rates, almost the same rankings can be produced. However, something that has only been mentioned in passing in most of the literature about the ACE studies is the nature of *reputational studies* and the methodological pitfalls of such endeavors.

The traditional reputational technique has been applied in sociological studies to the identification of community leaders. Charles Bonjean's studies in the early 1960's refined the method,[21] guided by the basic thesis that different groups with the community would perceive different individuals as leaders. James Preston concluded in a recent review of the literature that the basic *caveat*, when using the reputational study method, was that "the validity of the reputational technique *depends totally upon the selection of respondents*."[22] It seems a rather obvious statement, and yet when one is limited to the select group of "knowledgeable respondents" who are "best

---

[20]Cartter, p. 118.

[21]Charles Bonjean and David Olson, "Community Leadership: Directions of Research," *Administrative Science Quarterly* (December, 1964), pp. 278-300.

[22]James Preston, "The Search for Community Leaders: A Re-examination of the Reputational Technique," *Sociological Inquiry* (Winter, 1969), p. 40.

qualified" to judge the "profession," as in the case of the ACE reports, just such a stratified selection of "professionals" results.

The selection of respondents was similar in both ACE studies, but it makes more sense at this stage of analysis to turn most of our attention to the more recent Roose-Andersen research in 1970. One hundred and thirty deans of the largest graduate schools in the country were first selected. Of these deans, "the names of four scholars in a discipline were requested of institutions that awarded 2 per cent or more of the ten-year doctoral production in the discipline."[23] Two of the four were to be "senior scholars"; the third, the department chairman; and the fourth a "junior scholar." That was the selection process—four faculty members, including the chairman, all chosen by the deans of the largest schools, for each discipline.

> Nearly 8,100 names were obtained in this manner. Twenty-one per cent were identified as chairmen, 41 per cent as senior scholars, and 38 per cent as junior scholars.[24]

Of these, 6,098 returned useable questionnaires.

Despite the great amount of effort the studies expended to emphasize the scope of the survey, and therefore, presumably, its greater validity, the question still arises *whether this is an exercise in self-evaluation in which the graduates of the top-ranked schools make up such a large percentage of the respondents that they are essentially ranking themselves.* A look at some of the key disciplines in each of the major divisions will give some idea of how true a sampling one has for a valid reputational study.

The following tables indicate some of the major disciplines in each division, and give the number of departments being ranked, and *the percentage of respondents who did the rankings that received their degrees from the five top-rated departments.*[25]

---

[23] Roose-Andersen, *A Rating*, p. 28.

[24] Roose-Andersen, p. 31.

[25] These figures were taken from the computer printout information that resulted from the second ACE rating done by Andersen and Roose in 1971.

THE HUMANITIES

Disciplines	Number of Departments Being Rated	Percentage of Raters Graduated From Top Five Departments
Art History	30	55%
Classics	41	45%
English	92	36%
French	63	45%
Russian	17	59%
Spanish	65	29%
Philosophy	65	24%

In *art history*, 49 of the 99 raters graduated from the top four schools–New York University, Harvard, Columbia, and Princeton. In *classics*, Harvard was ranked first and Princeton third–between them, they accounted for 32 per cent of the sample of 116 respondents. In *English*, there were 239 raters of which 57, or 24 per cent, received degrees from two institutions–Yale and Harvard–which ranked first and second respectively. In *French*, the top three departments actually accounted for 34 per cent of the raters. In the small field of *Russian* there were 64 raters, and only seven departments made the top ranking, but here again the first three schools accounted for 55 per cent of the raters. In *Spanish*, of the 159 raters, 38, or 24 per cent, of them graduated from the top three schools. In *philosophy*, again Harvard and Yale accounted for 24 per cent of the raters.

SOCIAL SCIENCES

Disciplines	Number of Departments Being Rated	Percentage of Raters Graduated From Top Five Departments
Anthropology	42	47%
Economics	91	37%
Geography	34	34%
History	91	39%
Political Science	74	31%
Psychology	110	24%

Of the 126 raters in *anthropology*, 26 per cent of them graduated from the top three schools of Chicago, University of California at Berkeley, and Michigan. The top ten schools ranked and accounted for 70 per cent of the respondents. Chicago, ranked first, and Harvard, ranked fifth, supplied more than 30 per cent of the respondents. In *economics*, Harvard and California at Berkeley, ranked first and fifth respectively, account for 20 per cent of the raters. In *geography*, the top ten schools account for 20 per cent of the raters; the top three account for 25 per cent. In *history*, Harvard and Yale tie for first and between them account for 20 per cent of the raters. In *political science*, Yale is first and Chicago fourth, and together they account for 25 per cent of the raters.

NATURAL SCIENCES AND ENGINEERING

Disciplines	Number of Departments Being Rated	Percentage of Raters Graduated From Top Five Departments
Chemistry	125	28%
Geology	69	26%
Astronomy	35	40%
Physics	113	28%
Mathematics	102	27%
Botany	97	28%
Entomology	47	29%
Population Biology	68	32%
Mechanical Engineering	71	28%
Electrical Engineering	78	29%
Civil Engineering	65	31%

Of the 125 *chemistry* departments ranked, the top six account for 38 per cent of the raters. The top three of Harvard, California Technical and California at Berkeley account for 20 per cent. In *astronomy*, Princeton, ranked second, and Harvard, ranked fifth, account for 24 per cent of the raters. In *physics*, Harvard, California Technical and California at Berkeley tied for first and between them account for 17 per cent of the raters. In *botany*, California at Berkeley ranks first and Harvard and Michigan tie for

second. Together they account for 21 per cent of the raters. California at Berkeley ranks first in *entomology*, followed by Illinois and Cornell tied for second. These three departments, out of 47 ranked, account for 24 per cent of the raters. Finally, in *population biology*, of the 68 departments ranked, the top two, Michigan and California at Berkeley, account for 28 per cent of the respondents. In *civil engineering*, of the 65 departments ranked, the top seven account for 40 per cent of the raters, with the second ranked Illinois and the third ranked MIT together accounting for 20 per cent of the group. In *mechanical engineering*, again, the top seven departments out of 71 account for 41 per cent of the raters.

The data is solid and convincing. In the humanities, the top five departments in each discipline provide more than 40 per cent of the respondents. In the social sciences, the top five yield 35 per cent of the raters. And in the larger samples of natural sciences, the top five provide consistently between 25 and 30 per cent of the raters. If one fourth to two fifths of the respondents, who are the "dispassionate" judges of an average of 70 departments in each discipline, received their degrees from the top five institutions, *serious doubts must be entered as to any claims to a balanced and unbiased set of ratings*. Such statistics also show that *in some disciplines, close to half the departments being surveyed did not have a single representative alumnus among the raters*.[26] Cartter's general response to this built-in "weighting" in terms of size and "known prestige" schools is a classic elitist *non sequitur* which somehow again results in the "guaranteed judgment" of this group of raters, presumably because they are "distinguished" or have gone to "distinguished schools."

> The concentration of respondents at first suggests that this may be a source of biased rating in favor of the older prestige institutions. On the other hand, these are the departments which have turned out *the most distinguished scholars in the discipline, and thus their alumni are the individuals whose judgment should be particularly sound*.[27]

---

[26]Cartter, p. 101. Referring to the rating in political science, Cartter notes: "Thus 27 of the 65 departments surveyed did nto have a single alumnus respondent in our sample."

[27]Cartter, p. 101.

If we were to add to this inbreeding the additional percentage of respondents who did not do their degree work at the top five or ten schools, but who are now employed in these institutions, the bias of the sample would grow larger still. We can only conclude that the results ought to be treated with the skepticism that the biased sample seems to warrant.

Presentation of Results

Finally, it should be helpful to give an example set of tables from the Roose-Andersen study to show the way in which the results we have spoken of are presented. Andersen and Roose made one major improvement on the Cartter study, in keeping with their hesitancy about "ranking" quality in general. In place of the rankings within each category, they decided to group the schools into common categories, but then to list them alphabetically within each grouping. However, they admitted that the "dividing lines" between groups were themselves arbitrary selections.[28] They only ranked, then, the departments within the top category. For example, the groupings for Question A (quality of faculty) were made on the following basis:

Rating Term	Score Weight
Distinguished	5
Strong	4
Good	3
Adequate	2
Marginal	1
Not sufficient	0
Insufficient information	–

For Question B (effectiveness of program) the groupings were similar, but on a 0 to 3 scale.

Rating Term	Score Weight
Extremely attractive	3
Attractive	2
Acceptable	1
Not attractive	0
Insufficient information	–

[28] Roose-Andersen, p. 15.

For instance, for the ranking in philosophy, 19 departments received a score between 3.0 and 5.0 on the first question, and were ranked accordingly. In the next two groups—"adequate" and "marginal" or lower—they were simply listed in alphabetical order. In the second question the departments were ranked in the first two categories, 2.0 to 3.0, and 1.5 to 1.9. In the third category they were listed alphabetically. It is rather amusing that when Andersen and Roose dropped the titles "good" and "adequate" and substituted numbers, their apprehension about the evils of ranking seemed to be quieted.

There is one last set of observations to make about the methodology in the two ACE studies. Despite the continual admission by Roose and Andersen, and even by Cartter, that it is impossible to know on what basis the judgments were being made, or even what "quality" meant in the studies, both volumes are filled with statements that sound as if there is a solidly proven basis of judgment across disciplines and between different groups. For instance, Cartter, in the 1966 study, compared "graders" across disciplines as "lenient or tough," and then puzzled over the question of whether there is "more talent" in the disciplines that received higher average scores and "less talent" in the lower ones. For example:

> The hardest graders were in economics, mathematics, physics, and sociology. Economics is the only field in which approximately half of the departments reviewed appear in the two lowest groupings. In some fields two thirds or more of the departments were rated "Good" or better, but in the four fields named above and in history, political science, psychology, and geology, less than 45 per cent were in this category. *It is not clear whether there is in fact a more unequal distribution of academic talent in these fields*, or whether the nature of the subject makes possible a clearer separation of the good from the mediocre talent.[29]

If that kind of statement is not mystifying enough, given the almost total obscurity of what is being judged, Andersen and Roose spent a great amount of time comparing the two studies and "scores" given, in an even

[29]Cartter, p. 16.

more surprising way. They seemed to consider the methodological tool exact enough to compare the 1964 with the 1969 results, and talked in terms of higher scores meaning growth in quality during the period. They eulogized:

> The most dramatic development has been an improvement in the rated quality of faculty in a large number of graduate programs. Three quarters of the 1,600-plus surveyed in both studies *show increases in their "quality of graduate faculty" scores.* In 1964, 1,161, or 69.8 per cent of the rated faculties achieved the score category "Adequate plus" or better. By contrast, in 1969, 1,306, or 80.0 per cent, of the faculties included in both studies had equivalent scores.[30]

Andersen and Roose concluded that the number of quality departments had increased, since there was a greater overall percentage getting a higher score! This presupposes that the practically indefinable variables contained in the word "quality" remained the same and were understood in the same way; that the extent of the profession's knowledge remained unchanged although the overall number of programs had increased by 58 per cent from 1964 to 1969; and that a different group of raters would judge in the same fashion with a consistency great enough to compare the two studies.

---

[30]Roose-Andersen, p. 2.

V. THE USE AND ABUSE

OF STATISTICAL HEARSAY

It is practically impossible to delimit or measure the effects of the Cartter and Roose-Andersen reports. Clearly, certain audiences come into direct contact with the reports. The administrators and faculty of most institutions of higher education have seen them, discussed them, and some have made policy and personal decisions because of them. It is extremely difficult to know the extent of the ACE rankings' effects, but even a cursory exploration destroys the myth that "nobody takes these things seriously." Minutes from departmental meetings here, a speech by a Chancellor there, a letter of recommendation that floats to the surface, and the rankings' influence appears fairly extensive. One must always remember that there is still, ten years later, no good evaluative material, and so outdated statistical hearsay continues to hold forth. Donald Weismann, professor of art history at the University of Texas, describes this process exactly:

> Through the years I have found it somewhat terrifying that certain reports, summaries of this or that, opinion polls, and "studies" which were originally published and distributed with little or no effect are suddenly picked up and treated as "scientific," reliable, etc. instruments for measurement of this or that. Such reports and studies which on first appearance were obviously harmless or beside the point, suddenly gain unwarranted importance in times of some special stress. One of these stress conditions is with us now: how to economize within a burgeoning graduate education network. I fear that such things as the CAR Report may be used by university administrators and politicians as a measuring instrument. (Measure up or dissolve!) In my opinion it would be tragic for American graduate education if the CAR Reports and rankings were ever used to this end.[1]

It is still difficult to imagine how such crude evaluation based on

[1] Letters to Arrowsmith and Dolan, 1973. N.B. The abbreviation CAR appears in much of the correspondence as we begin to use it to refer to the combined Cartter and Andersen-Roose reports.

hearsay can be taken so seriously, but the story begins to have a familiar cycle of tolerance of a collection of hearsay that later turns into exactitude as a basis for policy decisions. At that same University of Texas, Dr. Roger Shattuck was chairman of Comparative Literature at the time of the Roose-Andersen survey. Although bothered by it, he "answered the questionnaire without much conviction and discounted the whole exercise as someone's innocuous make-work." Then he describes the too familiar process:

> Sometime in late 1970 or early '71 the new Roose-Andersen results were released with fanfares and eager attention from university administrators. Our graduate dean (University of Texas) convoked the graduate faculty and demonstrated, with slides of Roose-Andersen statistics and extrapolations, that we ranked among the six or seven top graduate schools in the country . . . the figures began to be quoted in university and foundation documents as if they came from an authoritative and objective evaluation of graduate programs in the United States.[2]

The transition from tolerated hearsay to scientific exactitude in the use of the ACE rankings parallels the same transition within the studies themselves.

Undergraduate Education *Not* Ranked

Some of the more subtle effects of such rankings take place because of what the studies were *not* concerned with, and never pretended to be. For instance, the rankings had to do with individual graduate departments, *not undergraduate education.* Yet, since the foundation of the studies was prestige, and size of the department had a great deal to do with the results, as we have noted, the overall effect was to strengthen large universities and emphasize graduate department strength over undergraduate priorities, since undergraduate excellence was not being proclaimed anywhere. Finally, as the media immediately began to add the cumulative rankings, and comment accordingly about the "great universities," the excellent undergraduate colleges were of

[2]Letters to Arrowsmith and Dolan, 1973.

course ignored, but the undergraduate priorities within the larger institutions were put under tremendous pressure. As a former Harbison Award winner for excellence in teaching observed about Notre Dame two years after the 1971 rankings had appeared:

> As recently as this fall I have seen a gimmicky set of priorities distributed by the English Department here with a copy of the Cartter rankings attached without explanation, like a prayer sheet or a fetish page. Notre Dame has had an adequate graduate school and has been and is a first-rate undergraduate institution. Nevertheless, in many departments including at times my own the graduate program is given primacy of attention without its having the substance to earn such attention.[3]

Dr. Allan Purvis, professor of English at the University of Illinois at Urbana, also commented about the pressures to play the prestige game of the graduate department at the expense of undergraduate instruction:

> In English there could easily be a move to the Orr model, but I am not sure that the English department would follow it. I must explain that there is a rift between the English department proper and that segment that is involved in teacher preparation: the powers in the department feel *the Cartter report is the only important fact in their lives. Undergraduate instruction is not important.* In sum, I do not think the college is ready to move in the way the Orr report suggests. I am sorry to have to say this, but there it is (italics added).[4]

These types of examples are spoken of frequently, but to find documentation is much more difficult. The following is an excerpt from a letter of the chairman of sociology at Vanderbilt to a professor of anthropology at Harvard:

> We are in the process of considering the promotion of X from associate professor to a professor of anthropology. I am

[3] Letters to Arrowsmith and Dolan, 1973.

[4] Letters to Arrowsmith and Dolan, 1973.

enclosing Professor X's vita, including a very brief statement of the program developmental work that he has done for us. Would you be good enough to evaluate Professor X's work for us? We could, if you desire, send you both his writings with which you are familiar.

My question can be sharply phrased: Does Dr. X's record justify professorship in high quality departments?

The Carter (sic) report uses phrases such as "distinguished," "strong," "good," or "adequate" to grade departments. Would Professor X's credentials warrant strong consideration for professorship in a department that on a *per capita* basis was considered good or better? (Italics added.)[5]

Since there appears no short-hand to aid this chairman in his quest for evaluation, and he appears to be fearful that some reaction to the content of what the man writes will prove too "soft" or unmanageable since it is obviously in an area outside his field, he uses the only categories at hand.

Faculty Evaluation Affected

The recruitment of faculty, especially within departments ranked among the top ten, is a process not unlike the trading of executives within a given industry. The names and work of those involved are well known, and the real questions involve how much grant money can Professor X bring with him or how many excellent graduate students, and so forth. Although the basic bargaining positions of such departments were probably affected in some cases, the overall results of the ACE reports in this area were probably minimal. But at the next level down, when faculty are being evaluated for hiring or promotion, the seemingly "objective" exactitude of "ranks" exerts its allure.

As we saw in the earlier discussion of the genesis of these reports, Cartter was concerned, as early as 1961 and 1962 with the oversupply of Ph.D.'s

[5]This letter is an example of the type of data that is extremely difficult to obtain but does begin to appear when word of such a study circulates.

and the need for future planning in higher education. The need for early and accurate data about manpower demands was clear and urgent. But the way the prestige rankings were eventually used probably bothered Alan Cartter, since he realized they were not the documentation necessary to make the hard decisions about growth or reduction of given departments. Of course, stronger advocates with less research concern began to use the rankings as surgical tools, even though they resembled rubber knives more than sharp scalpels. Dr. Salvatore Natoli, educational affairs director of the Association of American Geographers, noted that "we've received word that a few departments in prestigious institutions are undergoing rigorous administrator's scrutiny if they fail to fall in the top ten."[6] But for others, such as Dr. David Herbert Donald, director of the Institute of Southern History at Johns Hopkins University, the rankings were the perfect cutting edge:

> Since few universities these days can afford to do everything, it is highly desirable that their administrations use the CAR reports as a means of allocating funds and emphasis within the institution. Sometimes this means strengthening an already strong department and turning feeble departments into service departments; sometimes it means drastic rehabilitation and expansion of sub-standard but essential departments. I do not know whether foundations make grants on the basis of the Cartter reports, but they certainly should. Merit ought to be rewarded, not mediocrity. Finally, it is a very good thing that prospective graduate students are using the Cartter reports as a guide to selecting their universities. There is no sadder waste than sending a promising undergraduate for advanced training at Podunk University, where he will work under the inspired guidance of Professor John Doe and—if he is lucky—can hope at best for a job afterwards at Slippery Rock State. This country does not have enough brains to waste in such a falsely "democratic" educational program.[7]

The concluding example and its thrust are extremely close to the preliminary statements made by Alan Cartter early on in his study in 1964.

---

[6] Letters to Arrowsmith and Dolan, 1973.

[7] Letters to Arrowsmith and Dolan, 1973.

Resource Allocation Based on Rankings

Other administrators, however, found the use of the rankings a distasteful, but real, part of policy-making within their own institutions. In a letter Leo Treitler, a dean at Brandeis University, described his anger at the initial publication of the rankings. He was particularly offended by the low ranking of his discipline, and he and his colleagues attempted to defend themselves with the administration. He described the turn of events:

> My reaction was that we should try to cover this over as quickly as possible, for the low position that we had received was regarded by us as an insult. But we were one of very few departments that had been ranked among the upper twenty, and this has continued to give us credit within the administration. Now I am in the administration, and I can confirm that credentials bestowed through these rankings have a direct connection with the allocation of the university resources, *the last offense.*[8]

Another faculty member from a classics department in a Big Ten university reflected on the power that the facile numerical rankings exert in the absence of all other "quantified," and therefore "hard" and objective, data:

> I speak from personal experience on a retrenchment and reallocation committee. Our stated principles were to see how the mission of the school could best be served through the assignment of priorities for the use of dwindling resources. In the evaluation of departments ACE reports were to be regarded as one factor among many others, such as centrality to the teaching mission of the school and cost per credit hour. The committee members were at sea among these factors because they are difficult categories to grasp and to evaluate, but ironically the ACE reports appealed to some of my colleagues' more fascistic emotions which they identified glibly with a sense of high-level objectivity, evaluation by peers. Hence the ACE ratings contaminated all other ratings factors and became in fact the central rating

[8]Letters to Arrowsmith and Dolan, 1973.

factor in the recommended allocation of resources.[9]

The quantification aspect of the ACE studies makes the entire ranking system susceptible to the worst kind of summative and comparative action. Any university possessing three "adequate plus" departments and one "strong" obviously is *better* than a university with two "strong" and only one "adequate plus." Or is it? Even though we have seen some effects in struggles for resources within universities, sophisticated planners and policy makers in higher education would never be this simple minded. The president of the University of Nebraska published his presentation to the state legislature of the biennial budget of 1971 to 1973 in a bright orange and red pamphlet entitled *Case for Nebraska*. In the third section of that pamphlet, under the heading "How Good is the University Now?," President D.B. Varner told the legislature that the graduate programs at the University of Nebraska were "lagging." He stated his main proof in the following way:

> The Cartter report, a study of the quality of graduate programs published by the American Council on Education, gives an indication of the University's comparative standing in the graduate field. The report surveyed 29 graduate subject areas in major U.S. universities. The University of Nebraska had programs in all but two of the 29 (astronomy and classics). The report uses four ratings *in identifying clearly* acceptable quality: Distinguished, Strong, Good, and Adequate Plus. *Among the institutions in our Big Eight Conference, 16 disciplines at Colorado received ratings of Adequate Plus or higher; 15 at the University of Kansas; 13 at Iowa State; and seven at Nebraska.* And here is a further note: *all of the seven Nebraska ratings were Adequate Plus, the lowest on the scale.* This compares, for example, with Iowa State's ratings of two Strong, nine Good, and two at the lowest rank of Adequate Plus (italics added).[10]

Here is an example of cumulative, total institutional comparison, de-

---

[9]Letters to Arrowsmith and Dolan, 1973.

[10]D.B. Varner, *Case for Nebraska*, Biennial Budget Report, 1971-73 (pamphlet printed by the University of Nebraska), p. 7.

spite the fact that even in the studies themselves, the authors warn against the danger of this kind of game.

> Finally, no attempt has been made to aggregate scores or to identify "super," "distinguished," or "great" institutions. There are still serious difficulties in determining how disciplines should be weighed. . . . There is still the difficulty of evaluating and comparing institutions' functions and purposes.[11]

If we look at the recent history of the University of Texas, it's very clear that they began the "quantity-quality" syndrome in earnest fairly recently, and successfully by the ACE report standards. They moved from 15th ranked to 10th in classics from 1964 to 1969. In linguistics they changed from 11th to 2nd on the ACE rankings. It is interesting to read, then, what the president of the University of Texas, Stephen Spurr, said in his "State of the University" address for 1971:

> The best academic rating that we have of the University of Texas is from the 1969 study of graduate school programs of the American Council on Education. Although this rating does not deal with graduate professional programs nor with undergraduate education, *we may take it as approximately an evaluation of the University as a whole*, inasmuch as the rating of graduate programs in the traditional Liberal Arts fields are normally closely related to the overall caliber of the University (italics added).[12]

The ranking of the university as a whole from the ACE reports on individual departments was listed, but with the additional claim that—really—the undergraduate education was also implicitly included in the ranking. Then Spurr descended to a numbers game that made the ACE reports look like meaningless scorecards, and showed how they have been used at "prestige" universities.

---

[11] Roose-Andersen, *A Rating*, p. 33.

[12] Mimeographed copy of a speech given by Stephen Spurr at the University of Texas, September 30, 1971.

> Taking an average of twenty conventional liberal arts fields, the University of Texas ranks fourteenth. Among other state universities it ranks seventh. . . . *It ranks ahead of such distinguished universities as* the University of Washington, University of Minnesota, University of Pennsylvania. . . . *In the South, again considering the twenty conventional liberal arts fields, Texas as No. 14 is followed by* North Carolina (No. 21), Duke (No. 24) . . . and Vanderbilt (No. 50). This is good but it isn't good enough (italics added).[13]

Finally, in this vein, we cite a last example from Florida State University. In a newsletter, "Research in Review," the graduate dean, Robert Johnson, in May of 1971, wrote an article entitled, "On Building Quality in Graduate Education." He pointed out that Florida State University's chemistry department, in the 1971 ACE report, ranked 24th in the nation, "ahead of all chemistry faculties in the South below the Potomac except that at the University of Texas, which was 17th."[14] He summed up, a few pages further on:

> Despite the newness of graduate education at FSU, this institution was able to show up as ninth among institutions in states south of the Potomac in the number of disciplines given a ranking in the Roose-Andersen survey. The University of Texas led all of these institutions with 33 ranked programs, followed by the University of North Carolina with 25, Duke (23), Florida and Vanderbilt (22), Rice (19), Virginia (18), Tulane (16) and FSU (13).[15]

That was certainly lining up the pecking order in the South, according to Dean Johnson's interpretation. He then took the ACE results and his manipulation of them to go to his central point. "It is important," he wrote,

---

[13] Spurr, p. 1.

[14] Robert Johnson, "On Building Quality in Graduate Education," *Research in Review*, Vol. 2, No. 3 (May, 1971), p. 1.

[15] Johnson, p. 2.

"that Florida not spread available state and federal funds too thinly among many institutions, but, instead, maintain and build upon the graduate programs already receiving recognition . . . University of Florida and FSU."[16]

These indicate some of the crude, and basically foolish, uses the studies have been put to both within institutions and within regions—to "prove" excellence, greater prestige, or the need for additional funds from state legislatures. And this does not even touch on the uses made of the ACE reports by the Foundations and federal programs to award grants and thus to confirm the monetary pecking order.

ACE Reports Used to Guide Graduate Students

The second audience foreseen by Cartter and by Andersen and Roose was the graduate student. Sure enough, although a student rarely reads the reports directly, he or she "benefits" by their ranking many times over. A brief look at some of the more popular guides to graduate schools finds the ACE reports undergirding them all. In *The Random House Guide to Graduate Study in Arts and Sciences*, the reader is informed that the schools are not ranked, but are grouped into general quality categories, since reputation and quality are "always a little off." However, the reader is assured that the ACE study "has been incorporated in the data pertaining to individual departments." The results of the Cartter report are included by notation (one asterisk = extremely attractive, etc.) throughout the book. In the introduction, the structure and content of the book are explained.

> The second part of the guide is devoted to a discussion of the more common fields of study in which students are likely to be interested. A description of the usual undergraduate prerequisites for each subject area is followed by a list of departments offering study for the doctorate in each field, *including an appraisal of the quality of the major departments based on a recent study by the American Council on Education* (italics added).[17]

---

[16] Johnson, p. 3.

[17] E. Wasserman and Ellen Switzer, editors, *The Random House Guide to Graduate Study in the Arts and Sciences* (New York: Random House, 1967), p. xiv.

In *A Guide to Graduate Study*, edited by Robert Quick and published by the ACE in 1969, the preface naturally refers to Cartter's "comprehensive evaluation" of graduate education. Quick then follows with his commentary on the worth of the study.

> Evaluations by experts have their limitations. By their tendency to favor departments in older prestigious universities, for example, they exhibit a time lag in assessment of reputation. They do however . . . constitute what a court of law would call "*testimony of expert witnesses*" – those persons in each field who are well qualified to judge, *who by training are both knowledgeable and dispassionate*, who through professional activities are competent to assess professional standards, and who by their scholarly participation within their chosen fields have earned the respect of their colleages and peers.[18]

Now that is an impressive set of credentials for any student to ignore.

In *The Guide to American Graduate Schools*, published by Viking Press and edited by Livisey and Robbins, the preface mentions that there are few studies of quality in higher education, but the American Council on Education, "representing 1,400 colleges and universities," had published a study, "*acting as an agent for its member institutions.*" It concludes its discussion of the ACE study in this way:

> In spite of these reservations, the very substantial value of the survey cannot be denied. *The reader is urged to examine the report itself in its entirety. The brief notations in this book only suggest the barest outline of the scope and depth of the study.*[19]

Finally, in the somewhat less than exciting *Underground Guide to the College of Your Choice*, sandwiched between a discussion of the fluctuations

---

[18]Robert Quick, editor, *A Guide to Graduate Study* (Washington, D.C.: ACE, 1969), p. 6.

[19]H.B. Livisey and Gene Robbins, editors, *Guide to American Graduate Schools*, 2nd edition (Viking Press, 1967), p. xxxv.

of the price of "grass," and a not-too-adept description of the variety of sexual mores at Berkeley, comes this comment: "Berkeley has good undergraduate and excellent graduate departments for those who are on a formal learning trip. Its overall rating in graduate school standing is Number 1 by the American Council of Educational Surveys."[20] (The Cartter Report somehow sounds the same—even underground.)

The ACE reports also seem to have provided other of the disciplines, and perhaps the professional schools, a model for similar types of studies. They also seem to have animated the foundations to begin financing like proposals. The most recent example is a ranking of the departments of religion done by Claude Welch and financed by the Henry Luce Foundation.[21] Its methodology was not a reputational study, but the classifications, priorities, assumptions and basic agenda bear a marked resemblance to the ACE model. Actually the Cartter and Roose-Andersen rankings are used within the Welch report as a guide to "quality" universities, where it is recommended "quality" departments of religion should be supported. Alan Cartter was on the Advisory Board of that study.

Finally, *Change* magazine, in conjunction with Columbia University's Comparative Organization Research Program, appears to have hit upon "rankings" as its chosen genre, producing numerical rankings of professional schools and a list of "top" educators faster than the *Times* can tell us who the leading intellectuals are, or *Women's Wear Daily*, the sartorial misfits.[22] At any rate, the popularity contests are being played fast and furious in the educational media as well. If you join a parade, you must follow where it leads!

---

[20] Susan Berman, *The Underground Guide to the College of Your Choice* (New York: Signet Books, 1971), p. 56.

[21] See Appendix A.

[22] See Appendix B.

VI. SUMMARY

The Ace Rankings Themselves

As we have seen, the rankings of academic departments sponsored by the American Council on Education, although not unique, are considered by most of the research and discussion since 1964 as the seminal studies in this area. They established a methodology that is now being imitated. In addition, a compilation of the ACE rankings of 1964 and 1969 is providing the basis for categorizing institutions for the present research in the sociology of education. Most of the work, for instance, on faculty at "prestige" or "second and third class" universities, uses the ACE ranking as the base-line in forming the categories. So even a concern with the broader questions of assumptions behind judgments of quality—and with the effects of present-day evaluation on higher education—demands a careful look at the specific prototype studies that embody these assumptions and symbolize their national impact.

The object of the studies, as we have seen, greatly conditioned the choice of methodology and their ultimate form. There is growing proof in social science research that many of the critical questions are virtually answered in the choices of direction and methodology made before the formal project begins, and the ACE studies seem to provide a clear case in point. Personnel in the agencies that desired and financed the studies wanted a sorting mechanism based on criteria that were very clear in their own minds. There was little attempt to improve the state of evaluation, as the use of simplistic questions attests. The great weakness common to all reputational methodologies was accentuated in the ACE study. In community-power research, where reputational methodology originated, social scientists learned early that there was danger in identifying the leaders of *only* the social/economic strata of society that they questioned. Very different leaders were identified in a given community, depending upon whom one asked. From our analysis of the criteria of large departments, the deans associated with them, and the senior and junior scholars identified by these deans, the universe of the ACE studies was weighted in almost every case in favor of the same departments that emerged as the top five or ten. Knowing the methodology and its dangers, this should come as no surprise.

Cartter Report Identified Elite

We have mentioned that the preface of the original Cartter report was filled with rather blatant assumptions about an elite set of institutions and the need to identify them. Those assumptions cause no problem until they are applied to a study that purports to evaluate *all* graduate departments in the major academic disciplines. Then, what sounded like an entertaining exercise in popularity, becomes a so-called national evaluation on the basis of which legislators, presidents of universities and deans are urged to make tough planning decisions. Suddenly a narrow model of the university and its activities have been applied like a steam roller under the guise of two vague questions about "quality" faculty and doctoral programs. This focus on a certain minority of institutions and the single academic exercise of publication seriously distorts the reality of higher education, its faculty, its goals, and the needs of its clientele.

In addition, given the hearsay nature of the basic data of these studies, the rhetoric of the statistical approach is more than a little misleading. The transition from a crude, two-question reputational survey to the statistical correlation machinery of the later chapters is a rather contemporary rhetorical device. In the process, something highly questionable subtly takes on the air of hard, scientific research. The basic data of personal opinion is obviously not changed, but the impression is deliberately created that evaluation has risen to a new, quantifiable level. This movement from an exercise in popularity and prestige to a two decimal-place ranking is the essence of the ACE studies, and merits special attention. It is hardly something dark or sinister, but rather the clever use of rhetorical skills to persuade and impress. Through statistical manipulation, these studies appear to be something they are not, and that becomes very important when one understands that the game of national prestige is played for very high stakes.

There is proof that the uses of the ACE studies, although sometimes ludicrous, are quite obviously important in making policy and planning decisions in American higher education. There are enough examples of crude averaging of the results on the part of university presidents and chancellors to give clues to their use in the educational-political arena. It is even clearer that within many universities, the studies are powerful tools of influence for chairmen of departments and deans, whether they like it or not. Priorities are influenced, pecking orders established, and faculty chosen on the basis of

these ACE rankings, since there is little else to go on. And finally there is the more subtle, but even more serious, problem of the distortion of university priorities in favor of graduate over undergraduate education, and academic departments over interdisciplinary programs.

The sphere of influence of the studies continues to grow, as we have seen, providing a prototype for additional rankings in practically every academic and professional field. The rankings are being used in research in the sociology of education to form the categories of quality from "prestigious" to first, second, or third class. And the attempts to establish correlations between the rankings and other measures continues unabated. This is, of course, understandable since the relationship between prestige in the ACE studies and any criteria of academic quality remains open to question, unless we decide that quality equals quantity, and that we should be content with a rank based on size. This is easier to do, obviously, and provides much the same results as the ACE rankings. But the question of evaluation of *quality* in higher education, and its criteria, methods, and assumptions, is as problematical today as it has ever been. The ACE ratings did little more than cloud the issue with a prestigious sponsor and some statistical flourishes.

Evaluation as the Basic Question

Underlying the specific studies themselves is the more basic dilemma of a profession that is significantly divided about its goals and the practice which gives it its identity. A majority of the practicing academic professionals teach, and some of them excel at teaching. However, little if any training or evaluation takes place with respect to teaching, and professional prestige is *not* connected with it. A minority of academic professionals primarily do research and some excel at that. And research is what is evaluated for judgments of promotion and local and national prestige. This distortion not only of numbers, but of activities and goals as well, contributes to the confusion surrounding criteria for evaluating higher education. The control of the profession remains with the researchers and the research universities. The ACE rankings are a testimony to this. The resulting distortion has broad ramifications for the delivery of higher education, since it affects the practices and priorities of both individual professionals and their educational institutions.

This rather basic confusion within the profession makes it difficult to

build a strong evaluative base, but the problem is further complicated by the present mechanisms for evaluation. The regional accreditation process is being used for two totally different types of judgments. On the first level, it is used as a sorting mechanism for government agencies to determine minimal standards for qualifying for federal funds. This judgment has little to do with assessment of "quality" and more to do with judgments of good faith, fiscal responsibility, and some minimal educational standards. This practice will become increasingly problematical as the regional accrediting associations deny accreditation to institutions that are clearly more educationally valid than other post-secondary institutions which are "accredited" and judged eligible for federal funds by specialized accrediting associations. At the same time, the regional accrediting associations continue their on-going review process for *the improvement of quality* in higher education according to quite stereotyped criteria. But at least this function appears compatible with the idea of a private association whose member institutions purportedly join of their own volition. But the two evaluative functions are so clearly distinct and require such different approaches and structures that, unless the conflict is resolved, the confusion will continue to grow.

This leads to the critical question of the availability of information about "quality" education produced by the present evaluative structures. The list of accredited institutions ranges from huge research universities to tiny private colleges. They are lumped together without any specifics except that they are "accredited." When one adds to this list countless other categories of vocational-technical schools, the term "accredited" becomes virtually meaningless, and the student knows nothing more about what occurs at a given institution by knowing that it is "accredited." The only other source of information is the institution's own self-description in the marketing, and admittedly distorted, context of a catalogue. In short, the evaluative processes give the public no information about distinct learning environments and different settings or goals, let alone a disinterested third party review of the institution based on clearly stated criteria.

The historical growth of evaluation of higher education has constructed a many-layered/public-private maze of agencies that now finds itself back to square one. There is no clear set of goals emerging from the academic profession, since the professionals themselves are deeply divided over their own role and their primary activities. The government agencies, through the rapidly growing fiscal presence of student loans, are demanding a different sort

of evaluative process not unlike a Securities and Exchange Commission judgment of investment corporations. The public, with its growing awareness of the rights of the consumer to have correct information, and a clear need for diverse educational offerings, is exerting pressure for a third type of evaluative information. The new priorities will have to include the public's educational priorities as well as the professional's. Until the different needs are recognized, the confusion of relying on a single set of private associations to serve all the demands will continue. As a result, the present accreditation process will continue to provide fuzzy guidelines and no public information, and will continue to exercise its power to make judgments of eligibility for federal funds.

Future Types of Evaluation

The question then becomes one of future directions. Given the confusion of the academic profession about its own goals and practices, and the confusion of the broader evaluative mechanisms such as accreditation, the greatest mistake would be to close down examination into "quality" higher education. Yet that is precisely the effect of the ACE rankings and similar endeavors. By use of the rankings, all of the unquestioned assumptions about "excellence" emerge and are canonized: the single model full-university, the research professor, strict disciplinary orthodoxy, graduate education priorities. Instead of opening these questions to further evaluation, the rankings begin with these assumptions and simply extend their influence. Yet despite all these pressures for mediocre orthodoxy, the crying need for diverse quality in higher education remains. Some areas of improvement appear rather clear at this time.

First there is the *type* of knowledge we ought to be producing with respect to higher education. This involves the criteria that ought to be examined as well as the way those criteria ought to be measured. In our correspondence around this issue, Leo Treitler, Dean of the College at Brandeis University, spoke of an alternative:

> What I have in mind would be a system of rating, rather than a ranking, with a more inclusive set of categories in which each department evaluated can be rated on a fixed scale. Categories that come to mind: interest in and quality of teaching;

> importance of the published work of faculty; pursuit of unique or specially strongly cultivated fields; general competence in the discipline; placing of students in influential positions; leadership in the professions and professional societies, editorial boards, etc.; degree of adventuresomeness and flexibility in the department; etc. Then anyone using such ratings can weigh the several categories according to his own values.[1]

This recommendation does the one thing that educational evaluation is so reticent to do—set out the criteria clearly and urge public dissemination of the evaluative results to allow the consumer an informed choice based on personal values and goals. This type of qualitative description would appear to be the beginning condition for any significant change in the present closed structure.[2] It would allow for some differentiation of educational qualities and a modicum of informed choices based on this public information.

Once the decision has been made to open the exchange around what is really occurring in higher education, the two natural sources of sound information—the faculty and students—become pivotal. In our correspondence, responses in the direction of Dr. Frederick Crews, professor of English at Berkeley, were typical:

> I believe that a foundation could perform a valuable service by asking the members of each department to describe that department, the only one they are bound to know well. They could explain the purpose and philosophy of their graduate program, dwell on the particular advantages and handicaps, and characterize the sort of students they hope to attract. They might also

[1] Letters to Arrowsmith and Dolan, 1973.

[2] There was almost unanimous support for qualitative description as the most useful beginning tool. This observation by Dr. Robert Fuller, President of Oberlin College, was typical: "I do not think it is educationally important to rank departments numerically one to twenty. In the same spirit, I believe it is silly to rank twenty students numerically from one to twenty. It is much more valuable, and also less harmful, to *describe* the strengths and the weaknesses of various departments." Letters to Arrowsmith and Dolan, 1973.

> explain what they have been able to accomplish by way of redressing inequalities towards ethnic minorities and women. And finally they could be asked to show what typically happens to their students after they get their degrees. All this information would be valuable to prospective students and to foundations seeking to support special programs and causes.[3]

This excerpt continues the stress on clearly defined criteria, adding both social and career dimensions, but what is rather obvious–yet so rare–it calls for some self-description of a straightforward nature by the faculty themselves. This is the type of information, for instance, that is at present produced by accreditation self-studies, but it never sees the light of day. We now live with the academics' version of *caveat emptor*. The buyer bears the onus of the transaction, while the "professionals," offering their product/services, prefer to operate with as little light and information as possible on the transaction.

Finally, the student voice is growing in volume–especially as federal loan funds increase dramatically in post-secondary education. There are questions of honesty of advertisement and recruitment, not only in the area of specialized programs, but also in the traditional graduate and undergraduate offerings. Again, unless the educational institutions begin to produce better information about themselves and their results, the presence of the Federal Trade Commission (FTC), Department of Health, Education and Welfare (HEW), and other agencies is bound to increase. The avoidance of quality self evaluation on the part of the academics and their institutions may bring even tougher evaluation in areas that seriously distort the true goals of post-secondary education.

Some Future Directions

Despite the closing ranks of the academic profession and the strength of the present evaluative mechanisms, there appears to be a strong indication that some new directions are developing–including good information, publicly accessible, with clear and diverse criteria evaluated. Some further concrete steps might be:

[3]Letters to Arrowsmith and Dolan, 1973.

The present regional accreditation process must either decide that it is a federal tool making minimal judgments of standards for eligibility, or that it truly engaged in the active improvement of diverse quality in higher education.

In either case, the secretive and private nature of the process must open itself to public representation and share responsibility for public information about institutions of higher education.

Clearly public judgments of certification and licensing must take extreme care to separate themselves from dependence on private associations and/or solely the professionals involved, and must open that process to public needs and public wants.

A very worthy goal for foundations and federal educational agencies would be experiments in different ways of evaluating programs and institutions of higher education with strong emphasis on public dissemination of that information.

The central question of this report has been *evaluation*. It often appears to be a subject at the periphery of what occurs in higher education—an after-the-fact reflection on how it all seems to be going. I have tried to argue, on the contrary, that it lies at the heart of the matter. Until we recognize that what we try to measure shows our values and the goals we have preestablished for our educational institutions, we will be locked into only those goals. At present our evaluative processes measure a tiny and distorted portion of the possible criteria and goals we could and should be examining. The result is to narrow all alternatives to some slight modification on a single model of higher education. This happens because:

a. institutions of higher education prefer to operate with as little real measurement and public information as possible;

b. the academic profession is in a moment of closing its ranks and striving for respectability, so that it appears content to establish a kind of model that only 15 per cent or 20 per cent of the present professionals want to or can achieve;

c. the public, and particularly the student-client, until quite recently,

has been content to purchase with unquestioning acceptance whatever was delivered under the rubric "higher education."

The American Council on Education's rankings have become the symbol and the protector of this empty evaluative posture. The resultant distortion—based on the power politics of evaluation—has profoundly influenced the basic structure of higher education. There is still little or no training of Ph.D.'s in teaching skills. In fact, the professional reward system pays little heed to the entire area of teaching. The academic disciplines have grown so entrenched that significant interdisciplinary experiments are practically non-existent. If they do appear, most of the cost is borne by foundations or other exterior funding agencies, since the institutions wish to risk nothing in that direction, and the academics are thus assured of the lack of durability of the experiment. Graduate education, with its priorities and structures, continues to possess the strength within universities to determine content and delivery of undergraduate courses. Sensibility to the diversity of values and cultures in this society grows duller, since diversity is seen as less and less powerful in the hierarchy of prestige goals and mainstream recognition.

The ACE rankings are but a small part of all the evaluative processes, but they are also the most public, and they are clearly based on the narrow assumptions and elitist structures that so dominate the present direction of higher education in the United States. As long as our most prestigious source of information about post-secondary education is a vague popularity contest, the resultant ignorance will continue to provide a cover for the repetitious aping of a single model. The quickest way to achieve qualitative diversity is to clearly state some legitimate aims for higher education and to measure the present institutions against these criteria. We could then have an open and honest picture of successes and failures, with the opportunity to force both the academic professionals and the institutions themselves to generate some real alternatives. Until then, all the attempts to change higher education will ultimately be strangled by the "legitimate" evaluative processes that have already programmed a single set of responses from the start.

APPENDIX A

THE WELCH REPORT ON

GRADUATE EDUCATION IN RELIGION

The world of graduate education in religion is relatively small—69 institutions offer doctorates. However, this number also makes each of them fairly important when one considers that together they educate practically everyone who in turn trains the ministers, priests, and rabbis; write whatever scholarly work is produced in religion; and more and more teach the college courses in religion across the country. Therefore, when a 280-page report sponsored by the American Council of Learned Societies is published that purports to critically appraise and rank the graduate departments of religion, it is especially noteworthy. Claude Welch produced the study *Graduate Education in Religion: A Critical Appraisal* practically without the knowledge of the intellectual community at large or the departments of religion in particular. When *Time* magazine (October 18, 1971), however, informed the larger public of the study, and especially of its ranking, and the "marginal" and "inadequate" schools which Welch believed ought to be phased out were singled out, the report was taken in deadly earnest. It is of interest here precisely because it is greatly influenced by the American Council on Education's reports on quality of graduate education. The presuppositions are very similar, the rankings backed by a prestigious educational group look familiar, and, as we shall see, the Roose-Andersen rankings themselves are accepted as a crucial guide to Welch's judgments of quality in graduate education in religion.

The field of graduate education in theology had begun to change rather drastically after the Supreme Court decision in the Schempp case in 1963, which made the formal study of religion a legitimate endeavor in public institutions of education. Formerly the schools of theology, and the divinity schools, offered the major religious education, including the graduate level courses. During the 1960's a major requestioning of the structure took place and "clusters" of different denominational schools of theology began to form, and departments of religion began to flourish in the universities. It was to this set of changes that the Welch report, at least overtly, was a response:

> This study arose from the judgment that the clue to the future of these changes lay in graduate education: in the training of the scholars who are to be the teachers and researchers in the field.[1]

To understand exactly what the approach to this set of changes would be, it is necessary to note that Allan Cartter was on the Advisory Committee of the report and his influence is marked throughout. Cartter made his reputation in the early 1960's with his studies of manpower needs in education and made correct forecasts about the "over-supply" of Ph.D.'s at the present time. He also developed a strategy at that time to answer the problem. His counsel was to identify the true "quality" schools in each discipline, and cut back the training to the very best students at those institutions. That, in a nutshell, is the same argument that develops in the Welch report. Of course, the situation is made to order. The study of religion had hardly attained the proper status of an "accepted" academic discipline as yet, and there was a period of rapid growth in which one could find quite different goals and educational experiences at different institutions. In short, one could find some variety. If one is bent on defining "quality" in a distinct discipline that can be judged with common criteria and then ranked, then *different kinds* of graduate education are not the ideal subject matter. If you see the main problem as an over-production of trained "academicians," then growth and experimentation are more than suspect; they are out of place. However, Welch seems to commit the gravest fault when he applies the set of Cartter theses to the change from professional education in religion—which usually occurred in seminaries and resulted in degrees such as Doctor of Sacred Theology (S.T.D.) or Doctor of Hebrew Letters (D.H.L.)—to academic education. This latter goal is somewhat self-contained, and should take place only on a very limited scale, according to Welch, to respond only to the academic needs of the marketplace of religion departments in very selected prestigious universities. That, however, is the basic prescription of the Welch report.

Study Based on Visits and Interviews

As we have noted, the universe of doctoral programs that Welch had to

---

[1] Claude Welch, *Graduate Education in Religion: A Critical Appraisal* (University of Montana Press, 1971), p. viii.

study was small. There were 69 institutions in all, 50 of which were universities or other Ph.D.-granting institutions, and 19 were independent theological schools. Welch visited 59 of the institutions in one year, from November, 1969, to October, 1970, and interviewed deans, faculty members, and students. The statistical information Welch used was gathered from a series of questionnaires addressed to the same group. His goal, as he stated it, was to make a "critical evaluation" to better guide future decisions in the study of religion. The Henry Luce Foundation and the American Council of Learned Societies provided funds and official sponsorship for his study. The best description of what actually is being proposed in the Welch report is the outlining of the moves necessary to make the study of religion an "academic discipline." To achieve equal status within the present scheme of things in higher education, certain procedures must be established, and certain "kinds of training in certain places" must be sought.

Welch's basic message is that the study of religion must be separated from the professional schools that train clergy if it hopes to gain respectability. In Welch's opinion, unless it achieves this independence, there will be no real chance to become a bona fide academic discipline. Some of Welch's recommendations will establish the pattern:

> Since the academic study of religion at all levels is now extensively developed outside the areas of church-related schools and professional interests, *it needs more careful attention by academic associations* and in particular by public authorities.
>
> *Effective organs are required for representing* the academic study of religion to the several publics to which it is responsible, in particular to the civic community.
>
> Close coordination is needed between the several regional accrediting associations and the American Association of Theological Schools (AATS). . . . Together, *these agencies should discourage all non-university-related theological faculties from offering the Ph.D.*
>
> *The number of doctoral candidates in religion should be allowed to decline by at least one third, with more rigorous standards of admission being applied.*

The number of doctoral programs in religion and theology should be reduced significantly, though in a few cases new programs should be encouraged and supported.

Ph.D. programs should be encouraged *only in universities that are able to support superior graduate work in many areas.* The major burden of doctoral studies in religion and theology can and should be carried by such institutions.[2]

The strategy becomes fairly clear in this short course on how to gain academic respectability. Begin by lobbying with professional societies to get a share of the action; use the accrediting groups to apply the pressure to all programs not based at prestigious universities; cut down on both the number of programs and students, and be much more selective about both. Make sure the profession has a tight control on where this degree is offered and to whom. And finally get the supply back down where it's a seller's market. The consideration of other possible needs—for instance of an educated clergy—is completely ignored, while it is recommended that the places where they have traditionally been prepared be forced out of business.

One of the later and more intriguing arguments in the Welch report is the one based on need to change at the graduate level because of the changes taking place on the undergraduate level. Welch notes that more and more students are entering advanced studies in religion without the former "seminary" training, and that undergraduate training in religion itself has improved in quality, and has developed wider interests in different traditions, and so forth. He then begins a rather unique argument. First he notes that any graduate training "will require responsiveness at the graduate level to the changing interests and patterns in the undergraduate study of religion."[3]

One believes he hears professional training in an academic discipline trying to respond to needs and not making its own patterns and interests independently, and demanding these even of non-professional students at the undergraduate level. The following paragraph then comes as quite a surprise:

---

[2]Welch, pp. 7-8.

[3]Welch, p. 4.

> It is obvious that *some shortening* is desirable in time required *for the doctorate in religion,* since both in years of academic enrollment and in elapsed time, that period is now the longest of any field. *The only way to attain reasonable compression of the time span, without sacrifice of quality and of depth in specialization, is through more adequate integration of undergraduate with advanced studies* (italics added).[4]

Specialize earlier—that's the answer of how to get the graduate school over with more quickly. The undergraduate experience becomes the mirror of the professional structure, priorities, and skills. And thus one of the demands for academic correctness will have been achieved for the discipline of religion.

'The Making of a Discipline'

One often feels that the report could be subtitled, "The Making of an Academic Discipline," for each of the procedural steps for academic acceptance is outlined in detail. Surely one of the basic truths is that the only survival in a university setting today is as a *department,* and therefore, the only hope for survival is to enter making those sounds. When Welch asks himself the justification for separating religion off into a distinct "cognate area," in his terms, his answer comes to his real point.

> Is there anything that goes on in a department of religious studies that cannot take place in one or other of these cognate departments? *I believe not. The practical point is simply that in any university things grow only if they are tended, and that means being given an organizing center, a structure, a budget, some appointive control, and so on* (italics added).[5]

Once again, in viewing the formation of an academic discipline one can witness an additional concern for the profession's administrative clout inserted alongside the fear of over-supply, moving specialization up through the

---

[4]Welch, p. 4.

[5]Welch, p. 4.

undergraduate years, and designating the Ph.D. as the sole degree at carefully regulated institutions.

As Welch argues towards academic legitimacy for the study of religion, it is also interesting to notice how open the questioning is about intellectual quality, sorts of training, methods of procedure in graduate schools–especially when the discipline is in the unique position of having a rather open track with little former history to contend with. Welch states confidently:

> The kinds of questions to be asked generally about quality in doctoral programs have been extensively developed by previous studies of graduate education, by accrediting agencies, and by such groups as the Council of Graduate Schools in the United States. We need not duplicate that work.[6]

He then goes on to mention some of those standard items, such as strictness of admissions by intellectual "merit," small programs based on the "need" of the profession, the prestigious degree of the Ph.D. with other inferior or different paths or degrees excluded, full-time students preferable–especially those right out of specialized undergraduate programs. It seems unnecessary to spend further time proving that Welch is fairly uncritical about the general premises of graduate education.

When Welch speaks of the real credentialing that he sees taking place in higher education, one can understand much better how his report ought to be viewed.

> Accountability to the academic community, which is not simply separable from the responsibility to the civic and religious communities, though we must focus on it as our primary concern, is less a matter of formal accreditation processes than of *the informal credentialing* that goes on through *recognition of the leadership of certain institutions and that approval or disapproval that is expressed by comparing the work of other schools with these* (italics added).[7]

---

[6] Welch, p. 66.

[7] Welch, p. 86.

The basic process then, which this report is meant to fulfill, is to define, or accept in Welch's case, a model of excellence, compare various institutions to it, publicize the results through a national group, and so form "criteria for decision-making" for everything from federal expenditures to graduate student choices. The result is a sameness in which each new discipline is carefully drawn into the same circle of universities to produce the same kind of "scholar" in the same way.

When Welch summarizes his categories for rating the doctoral programs, he mentions three sorts of criteria.

1. First, there is the sort he calls more "objective," a la Cartter reports, such as library holdings, selective admissions, proportion of full-time to part-time student, and amount of financial aid.

2. Second, he makes what he admits is a much more subjective judgment of faculty quality based on his own "extensive" visitation of institutions.

3. But the last category deserves to be quoted in full:

> Finally, I have taken into important account the actual or potential relation to a university of the first or second rank. For this purpose, I have *drawn particularly, on the 1969 Roose and Andersen ratings of graduate programs (1970),* specifically the ratings of the following departments in the humanities and social sciences: classics, English, philosophy, anthropology, history, and sociology. *Since it is not possible to carry on doctoral studies in religion of the highest quality* except in relation to a strong university, *this consideration is of special significance* (italics added).[8]

For Welch then, as for many others, the Roose-Andersen report has without further question established the accepted "quality graduate programs." One should also note that it is forthrightly declared that no other "quality"

---

[8]Welch, p. 89.

education in any field can now take place anywhere save at these institutions. The one model, the one elite circle, the one pattern of training and priorities is uncritically accepted. The goal is to bring the study of religion along in the *snake*-like parade.

Religion Programs Ranked by Welch

In the ranking, under the first heading of "older and established programs of the first rank," Welch lists Chicago, Columbia, Duke, Harvard, Princeton Seminary, Union Seminary, Vanderbilt, and Yale. Of these top eight schools, six of them were represented on the thirteen-man Advisory Board of the Welch report. In the process of ranking the programs into six divisions, Welch is very self-assured in his opinions about what ought to happen at each institution, based from all indications on, at the most, two or three days spent at each site. His overbearing "confidence" about where an admittedly confusing field of study is, and where each of its departments and schools should proceed, is staggering. His comment on the religion program at Drew University is typical of many comments:

> Drew once possessed a strong theological faculty and a dynamic doctoral program, but since the resignation of a large number of that faculty in the late 1960's, it is doubtful whether it has been wise for the university to re-enter doctoral studies in religion.[9]

Or:

> Both Fordham and Temple have expanded beyond all reason and with low selectivity. . . . Temple, with a core of good faculty, has been quite undisciplined and unfocused both in areas of study and in programs; serious efforts toward discipline and focus have been undertaken since 1970, but supporting resources in collateral departments and especially library have been poor.[10]

---

[9]Welch, p. 94.

[10]Welch, p. 95.

This may be the kind of advice someone would ask of a consultant who had spent a certain amount of time at a school, and who had been invited to do the study, but not from a flying visit from a self-appointed organizer of the discipline. Even for Welch's top eight institutions, in his terms the "oligopoly or grandfather" institutions, he is generous with his advice:

> Chicago needs to be more concerned about the attrition rate, Columbia and Princeton Seminary about the high proportion of students writing dissertations *in absentia*. While Columbia has major strengths in the Union Seminary faculty and in collateral departments in the university, the faculty resources in the department need to be expanded.[11]

So the divisions are basically based on the ACE ranking and Welch's opinion of the faculty after a few days visit. Still, after all is said, one wonders what has really been ranked, as the large categorization by size of program (faculty) and prestige of the university tends to completely neglect the specific areas of competence within each school or department, where the true "quality" of resources lies. It is amazing to hear Welch deal with this problem. As he begins his rankings, he expresses this caution. *"This summary is not organized in relation to specific fields of specialization where the final test of responsibility and adequacy lie."*[12] Then, as he concludes his rankings and off-the-cuff advice to each and every school, he qualifies his ranking in the following way:

> First, the classifications have not been made in relation to *special fields* of study and therefore *conceal important differences in the field offered and in the quality of work*. These general estimates must therefore all be viewed in relation to the spectrum *of subfields actually offered and to specific strengths and weaknesses. In a review by special fields, particularly if made without relation to the age or size of the programs, quite different groupings would emerge* (italics added).[13]

---

[11] Welch, p. 91.

[12] Welch, p. 88.

[13] Welch, p. 96.

One must review this last set of remarks carefully.

1. The "real test of adequacy," in Welch's own terms, lies at the level of specific subfields.

2. *If* one did *not* use the criteria of size and age (which Welch admittedly does), but the criteria of strength and weakness of special subfields;

3. *then, "quite different groupings would emerge."*

We then have a general ranking—based on size, age, and general institutional prestige—that tells us nothing about the specific subfields where Welch admits the true quality lies.

What then do we have? Who really wins or loses, and on the basis of what criteria, related to what kind of quality in graduate education in religion? The answer, I'm afraid, is that we have an opinionated *tour de force*, based unquestioningly on the ACE reports' priorities and assumptions about "quality" education; a direct use of the combined ranking in social sciences and humanities from the same reports to establish the only institutions where quality religion departments *could* exist; and finally, such a general set of opinions that, by Welch's own admission, the true level of quality in religious education is missed entirely.

APPENDIX B

THE RANK AND FILE

Apart from some of the more bizarre examples of "rankings," the endless process of reputational studies appears not only not to be diminishing, but if anything, gaining in frequency. And the more the genre progresses the closer it edges to a parody of itself. In the Fall, 1972, issue of *Public Interest,* a social science quarterly, a list of the 70 most prestigious contemporary American intellectuals appeared. The methodology used is a nice, if humorous, example of the kind of process described in the ACE studies, where people are asked to rank themselves as important. They began with a list of 22 leading American intellectual journals, and we have made some comments in the body of the text that would make one interested in the makeup of the editorial boards of those journals. It was then determined that some 8,000 persons had written for those journals during a certain period in the late 60's. In some sort of frequency test, these 8,000 were cut to 200 finalists. Then of these 200, 100 were interviewed and each was asked to name which on the list of 200 had the highest prestige or reputation with the person being interviewed, and his or her intellectual circle. The answers were tabulated and the result was the list of 70. The two editors of the magazine, *Public Interest,* are Daniel Bell and Irving Kristol. Interestingly enough, Bell was ranked within the top ten and Kristol in the second ten. And although the list is perhaps less interesting than the ranking of the ten worst dressed women—and perhaps has even less impact—the process of beginning with journals and publications and proceding to ask the people who are to be ranked to rank themselves according to "prestige," defines what is now becoming a fairly familiar taxonomy.

One can have fun, as the *New York Times* did, with such a ranking, with the knowledge that very little has been affected by the process or the publicity.

Rankings Are Likely to Continue

Prestige and reputation, of course, are facts of life, and *Harper's* will

continue to publish its lists of 20 of the best-but-unknown hotels in the world; *Town and Country* will continue to rate the top 50 prep schools. Each, of course, will have some impact on those ranked. The schools will probably have a few more applicants; the hotels will probably become crowded and get worse. But a pattern is emerging of a more serious nature. For the rankings based on prestige, with all the pitfalls that have been described, are beginning to appear more and more frequently as the acceptable and unquestioned procedure without much reflection upon the effects or limitations of such endeavors.

MBA magazine in December of 1974 canvassed its clientele in business schools of the country and produced a double ranking—one on academic quality and one on the monetary employment value of the "leading" programs in graduate business education. The set of two questions is always interesting, for as we saw in the Cartter and Roose-Andersen reports, the first question had to do with quality of the faculty and the second had to do with the quality of the doctoral program—an interesting dichotomy in itself. In the *MBA* rankings the first question had to do with academic quality, but the second had to do with the value of the degree in terms of job and earning power. At least some of the more implicit values underpinning prestige judgments were being dealt with more explicitly.

But closer to our concerns, both because there are some linkages philosophically and methodologically to the ACE rankings, and because there are questions of breadth and scope of publicity similar to the ACE studies, are the rankings of professional schools published by *Change* magazine in the last two years. In November, 1973, *Change* published an article written by R.Z. Margulies and Peter Blau. Blau is director of the Comparative Organizational Research Program at Columbia University and Rebecca Margulies was a doctoral student in sociology at Columbia and a member of the same research program. The lead paragraph of the article sets the context in which—in the authors' minds—the ranking of professional schools is important. They mention the increase in the labor force of college and professionally trained individuals and stress the point that for the first time in 20 years, trained professionals in several fields outnumber the jobs available. But the role, in their minds, of prestige and its oversupply is rather interesting: "Since holders of masters and doctoral degrees have proliferated in the labor market, where one has come from, rather than the degree itself, may represent an

increasingly powerful passport to entry into professions."[1]

Study Linked to ACE Rankings

The most interesting linkage to the ACE rankings of 1966 and 1971 is the way they are viewed as a prototype for this kind of study. It seems that once a model is accepted, its underlying assumptions cease to be questioned. It is taken as the established piece of research or literature from which new endeavors simply build. When the authors of "America's Leading Professional Schools" asked the question about academic quality and the criterion for this quality, their facile answer was rather interesting: "The criterion that has been found to be most reliable is the judgment of experts in the field." The authors then immediately made reference to Cartter's study of graduate education and the follow-up study in 1970 by Kenneth Roose and Charles Andersen. They defined their experts in ways quite similar to definitions in the ACE studies.

> We used a similar approach [to that of the ACE rankings], except that we asked 1,180 deans of professional schools, not a sample of faculty members, to name the five most outstanding schools in their own profession. We assumed that deans as the leading officers of each school, have knowledge and experience for these evaluations.[2]

Another observation in terms of similarity of backgrounds and linkages of a longer-term sort is that these studies come out of the Comparative Organizational Research Program at Columbia University under, interestingly enough, a grant from the National Science Foundation. Here is the same interest by the same groups which appeared practically 15 years earlier at the origins of the first Cartter-ACE report.

This study by Blau and Margulies encompassed American professional schools that fall under a double criteria. They must be first of all accredited

---

[1] *Change*, Vol. 5, No. 9 (November, 1973), p. 21.

[2] *Change* (November, 1973), p. 21.

and secondly university affiliated. That leads to, especially in the professional area, some rather interesting contradictions. A few of the top ranking schools in professional areas were not among the schools that were sent questionnaires, either because they were not university affiliated (i.e. free standing schools of business and so forth) or they were not properly accredited even though top ranked. The authors warned:

> Thus, we have five top ranking optometry schools, although questionnaires were sent to only four schools since only these four are both accredited and university affiliated. The School of Optometry at Indiana University for example does not meet our criteria for accreditation. Other schools which were not sent questionnaires such as Graduate Theological Union do not meet our second criteria, university affiliation.[3]

So, because of those unquestioned categories of accreditation which we saw were the backdrop of the ACE studies, the School of Optometry at Indiana—although ranked first in the country—is not accredited and therefore did not receive a questionnaire. The Graduate Theological Union at Berkeley—ranked fifth among professional theological schools—did not receive one either.

High Prestige Equals Less Interest in Prestige

Another rather interesting aspect of the methodology was the distribution of responses across the different professions. When one is only questioning deans of accredited and university affiliated professional schools, there is a rather limited universe to begin with. So the researchers broke their tables into two parts: the first category, in which they had more than 20 responses and felt that these were more reliable, and the second set of ratings for other professional schools, where ratings were available from only a small number of deans. In several of the professions, although the number of deans defined in the universe was large (for example, medicine or dentistry), the responses from those deans, even after a second questionnaire, was minimal. The

---

[3] *Change* (November, 1973), p. 22.

authors made a rather interesting conjecture about the causes of this lack of response:

> A look at the response rates shows that one explanation for the differences between types of schools is the prestige of the profession. The largest proportion of deans who did not respond are in those professions that research has shown to have highest prestige, the extreme example being medicine.[4]

This represents a rather interesting application of reverse english, since such reputational studies are purportedly designed to be measuring prestige of institutions. But in the opinion of the researchers performing this one, the higher the prestige of the profession, the less interested they are in rankings of prestige. The deans may, of course, consider such rankings nonsensical or useless for other reasons, but when researchers are involved in a study of prestige, perhaps they begin to think in terms only of prestige.

As validation for the results of their survey, Blau and Margulies referred triumphantly to their correlation with the results of the Cartter study ten years earlier in one specific case.

> Engineering is the only field included in our research that was also studied in Cartter's earlier study of graduate schools. Although our ratings refer to the entire school and his, ten years earlier, are based on ratings of the four major engineering departments (chemical, civil, electrical and mechanical), our ratings and his summary ratings for engineering schools are in close agreement. His five top engineering schools are the same as ours, except that in our case a sixth, Michigan, is tied for fifth place.[5]

Not only is there this growing proof that the ACE rankings are considered the prototype of this type of study, but that they are beginning to be referred to as the baseline data on academic reputational quality. It is against this data that present studies are being compared.

---

[4] *Change* (November, 1973), p. 22.

[5] *Change* (November, 1973), p. 22.

By the time *Change* had gone to press with their winter's 1974-75 issue, Peter Blau and Rebecca Margulies were back with the second installment of their rankings entitled "The Reputations of American Professional Schools." The article was in some sense occasioned by the reaction to their article a year earlier. As the authors referred to the reaction, they mentioned, "Some of it turned out to be highly critical of our inquiry, and we felt professionally compelled to review our research and replicate part of it with a wider data base."[6]

Broad Criticism Refuted by Authors

It is instructive to watch the authors refute the broadest kind of criticism that opposes all rankings. Their interpretation of the motivation behind this opposition is that the comparisons hurt many good schools that may not be at the very top of their field, and this is especially dangerous in a period of limited resources. Once again the assumption that there is a single continuum of quality from the top to the bottom underlies the interpretation. There is no recognition of the fact that quite possibly professional schools with quite different missions and diverse goals could and should be possible, so that quality would mean quite different things in each of the diverse categories.

Their general defense of the rankings focuses primarily on the right of the public, especially prospective students and employees, to have a knowledge of how good or bad schools are. It appears though, that the question of whether or not rankings do provide a set of just and reliable public information is begged. The case has been made within this essay that rankings *de facto* are extremely damaging to any kind of diversity of program or institution, and are based upon such unexpressed judgments within the profession that they entirely distort whatever values the *consumers especially* might be interested in. For instance, the most important consumer, even in the case of the article by Blau and Margulies, was the prospective student. However, there has been proof time and again that many of the "top departments by reputation" (and this applies equally to professional schools, it would appear)

---

[6] *Change*, Vol. 6, No. 10 (December-January, 1974-75), p. 42.

are vicious and inhuman environments for students, and sometimes they appear to be the antithesis of a good and creative learning environment in which contact with professors is solid and frequent, work relationships with fellow students are productive and trustworthy, and the general rewards are based on acquisition of knowledge rather than on parroting of pet theories and on competition for grades. It would appear then that the "consumers" are very *poorly* served by such rankings which mirror, at best, somewhat informed opinion about the existence of a research institute atmosphere and, at worst, numerically classify schools and programs where students are last in a long list of priorities, and where the star profession researchers, their interests and prerogatives, stand at the head of the list.

Against the criticism that reputational rankings are in no way an objective measure of quality in education, Blau and Margulies responded with a paragraph that is well worth examining more closely:

> A shortcut for measuring professional quality is measuring professional reputation among experts. Social science research has shown that the best judges of the quality of the work of a professional remain his or her colleagues, and we must assume that the same is true for judging the quality of professional schools. The central position of deans and their responsibility for recruitment should make them particularly well informed about the quality of the schools in their field. To be sure, the resulting ratings of school reputations are merely the opinions of these experts, but all professional reputations are no more than the opinions of fellow experts.[7]

The first sentence is once again the disclaimer that at the same time subtly sets up an equation that has yet to be proved, and is the critical one in any kind of reputational study. At the same time that the authors admitted taking a shortcut by way of the rankings, there was the immediate association that measuring professional quality is the same as measuring professional reputation among experts. That is an extremely dubious assumption for all the various reasons which we have examined in this essay. There is the

---

[7] *Change* (December-January, 1974-75), p. 42.

difficulty of the self-selecting universe of respondents, the totally undefined makeup of "quality" in the questions, the complexity of sub-divisions within an academic department which surely is magnified when one is speaking of a large professional school, and so forth.

As for the second point in the paragraph, it is very clear that peer evaluation works reasonably well for evaluation of colleagues for tenure, for the evaluation for research bids, and so forth. But this is based upon first-hand judgment and contact with the person and/or his or her work. This type of judgment is *not* a judgment of reputation, and most professionals would be extremely reticent to make these kinds of judgments on only the reputation of a colleague whom they have never met or whose works they have not carefully examined. Therefore, once again judgments of quality are not equal to judgments of reputation.

Social science research has in addition shown us that reputational research is highly suspect, and extremely vulnerable to distorted results based upon the selection of the respondent universe. As for the second assumption in the sentence, that "the same is true for judging the quality of professional schools," the assumption is wrong on two counts--first, because the first part of the statement does not hold up, and second, because it would only be in the wildest stretch of the imagination that one could argue that the same kind of knowledge is possessed by the individual professional when he or she passes judgment on the quality of a colleague's work in the specific cases we have mentioned above, and when they rank schools. It is not a first-hand knowledge; it is not knowledge of a specific work of specific individuals; and the distance between the kind of specific judgment made in one case and guesses based on hearsay in the other case can hardly be considered comparable.

As for the validation of deans as the most informed group in terms of quality of schools in their field, in the case of major academic departments and major professional schools, their usual role is to appoint a search committee from within the proper discipline to fill a given slot. They are ordinarily not, nor do they pretend to be, authorities within that field, and they usually accept the recommendations of those committees as being by far the best informed in the specific areas in which they are recruiting. Academic excellence lies within the subdivisions of professional schools and academic departments and that is where the judgments are appropriately made. Deans can

ordinarily speak with knowledge in their own academic area, as can other faculty, but no further.

Rankings Based on Vicious Circle

Finally, the last sentence perhaps says more about the vicious circle within which the reputational rankings turn than anything that could be added. As the authors point out, ratings of reputations are opinions of "experts," and all professional reputations are no more than the opinions of these experts. That, of course, is precisely the point. That circle is a reputational circle; it has nothing to do with academic excellence and quality unless one is willing and able to begin to talk about the makeup of quality or academic excellence, and to define it, especially in terms of specific functions or particular missions.

Funds to complete this second round of reputational studies about professional schools were provided by the Russell Sage Foundation, and again by the National Science Foundation. In order to increase the response rate in the second study, the researchers asked the deans only a single question, and probably the most simple one that can be used in a reputational study: "What in your opinion are the top five schools in your profession?" There was no mention of academic quality or faculty excellence or doctoral program quality. The reasons for the simplicity was to accommodate busy deans and to increase the "response rate, thereby increasing the reliability of the rankings." Somehow what was being asked had become less important than a large sample.

The speed with which these studies fold in upon themselves seeking collaboration is alarming:

> There is a close correspondence between the reputation ratings obtained in the new study and the old study published in *Change* last year, and this fact justifies our original assumption that the earlier ratings were fairly reliable in spite of the low response rate.[8]

---

[8] *Change* (December-January, 1974-75), p. 43.

The authors were able to find two rankings done recently along quite similar lines. Again at Columbia a recent survey of 580 faculty at medical schools produced rankings of the top five medical schools corresponding to those of the deans. Even broader, and for the authors, stronger, corroborative evidence exists in a 1970 rating of programs in librarianship. About 150 respondents ranked three things: doctoral programs, masters programs, and quality of faculties. They found programs highly rated in one area tended to be highly rated in all. The authors feel this is a ranking "on the basis of specific criteria" which corresponds to their own rankings, "that fail to specify any criteria," and shows the value and validity of theirs. Specificity is a relative thing, but the questions were similar to the ACE double question, and the lack of concrete information in any of the areas could produce similar results in all.

An attempt was made by the authors to establish a correlation between high rank and other variables, such as the size of budget, separate professional library, and so forth. There seems to be little consistency, with separate library correlating in some professional schools (law, business, education), and hardly any at all in engineering, pharmacy, social work, and nursing. Other discrepancies appear when looking at total budget, percent of professional school budget from university, and so on. The ideal correlation in the authors' minds, of course, would be with the 1970 ACE rankings used as the gauge of the total university's prestige. The authors do find that the "top" universities do have the greatest number of high ranked professional schools, but the correlation is less than impressive.

> Nevertheless, the data on all professional schools do not support this expectation. The reputation of a professional school and that of its university (based on the Roose-Andersen ratings) are only weakly correlated, as row 4 in Table 2 shows, and these correlations are probably spurious.[9]

It is clear that rankings by reputation on the ACE model are not only continuing but growing. Each builds upon the other and defends the process by reference to those done previously. Each time there is less concern expressed for the effect or the questionableness of the process. In Cartter's first report there was at least careful and self-conscious awareness of the

---

[9] *Change* (December-January, 1974-75), p. 47.

dangers of rankings and choice of respondents. Ten years later, *Change*'s articles were a mixture of self-righteousness ("consumers ought to know") and only a bit of doubt at the difficulty of what they are claiming to measure. As they admit:

> Furthermore, different professional schools concentrate on different endeavors . . . and weighing the various programs to arrive at an overall index of quality would be quite arbitrary.[10]

That, of course, is the central question. There seems little awareness of what effect the ranking steamroller has on the diversity they concede is present, and that *de facto* makes what they are attempting impossible.

Finally a Ranking of Educational "Leaders"

Once on track, *Change* magazine seems captured by its own ranking frenzy. By February, 1975, a "Who's Who in Higher Education" article appeared with the promise of a repeat in 1976, assuredly to avoid "freezing" the list! Some 4,000 college presidents, foundation executives, journalists and bureaucrats were quizzed, and 616 individuals were nominated. *Change* published a list of the 44 most frequently mentioned. In the article there was not much self-consciousness about how the respondents were chosen. They noted, apparently with some chagrin, that only four women were among the "top forty-four," and no blacks. It would be interesting to know what percentage of their respondent group was black and/or female.

More interesting, however, is the apparent lack of awareness of what kind of a process the listing of 44 people in "Who's Who" is. What good is it? What does it represent? What are we supposed to do with it? There is some bit of *Change* analysis that revolves around a profile of "leadership," that indicates that this is the justification for the list. But there is very little realization that in reputational studies you get what you ask for. No young "leaders" appear on the list. How many young were questioned? Few women? Any correlation to the number of female respondents? Was the list dominated by white males? What was the universe of those asked?

---

[10] *Change* (December-January, 1974-75), p. 42.

Along these lines, *Change* seemed taken aback by the limited perspective of their respondents.

> A very large fraction of responses reveal an intense parochialism. Hundreds of names were forwarded whose fame could hardly extend beyond a single campus or the county line. Some college presidents felt compelled to list their ten top-ranking faculty and vice presidents. . . .[11]

The message may be two-fold and a paradox. *Change*'s concept of "fame" may have little to do with *leadership*, and the presidents could be telling them this in so many words. And at the same time it may be perfectly appropriate for each college and university in the country to have their own set of leaders who "contribute most significantly to the thoughts and actions of American higher learning." That may be true scope and diversity.

In short, *Change* may be looking for, and in every sense creating, the kind of "leadership" it assumes ought to be present. When Gerald Ford, Sid Marland, and the editor of one's own magazine appear on the list of 44, along with the heads of ACE, AAC, AACJC, NASULGC, AASCU, one begins to wonder what is being served up. There is a moment of doubt in which the imagination has visions of a *National Lampoon* article mistakenly inserted by the printer. But the serious analysis continues, although there is a moment in which the editors of *Change* do suspect that what they are measuring may not really exist. There is more than a little doubt in their last conjecture.

> . . . But barring an extraordinary windfall, a more vigorous national leadership in higher education is unlikely to emerge. What is perhaps more likely is the regeneration of local and regional initiatives of extraordinary effectiveness that could spark emulation elsewhere.[12]

---

[11] *Change*, Vol. 7, No. 1 (February, 1975), p. 26.

[12] *Change* (February, 1975), p. 26.

Their second set of instincts may be correct. The South may not have any leaders on their list of 44, yet men and women are influencing countless institutions there, from Texas to Tulane to New College in Sarasota. Sixty-six per cent of their list came from the East, but 66 per cent of those having an impact on education do not. It is the kind of distortion that proves nothing. Clark Kerr, the biggest vote getter, was mentioned by only 25 per cent of the voters, Reverend Theodore Hesburgh by 15 per cent, and John Brademas and Roger Heyns by 9 per cent. This less than overwhelming endorsement made up "the top four." Presumably there are some on the list of 44 that were mentioned by 1 per cent or 2 per cent of the voters (college presidents, etc.). *Change*'s "Who's Who," may not really be out there, and that may be a very good thing.

BIBLIOGRAPHY

PRIMARY SOURCES

Cartter, Alan M. *An Assessment of Quality in Graduate Education.* Washington, D.C.: American Council on Education, 1966.

Roose, R.D. and Charles Andersen. *A Rating of Graduate Programs.* Washington, D.C.: American Council on Education, 1970.

Welch, Claude. *Graduate Education in Religion: A Critical Appraisal.* University of Montana Press, 1971.

SECONDARY SOURCES

Berelson, Bernard. *Graduate Education in the United States.* New York: McGraw-Hill, 1960.

Bevan, W. "Higher Education in the 1970's: A Once and Future Thing," *American Psychologist,* No. 26 (Winter, 1971), pp. 537-52.

Bonjean, Charles M. and David M. Olson. "Community Leadership: Directions of Research," *Administrative Science Quarterly,* No. 9 (December, 1964), pp. 278-300.

Eells, Walter C. "Leading American Graduate Schools, 1948-1958," *Liberal Education* (March, 1960), pp. 16-20.

Elton, Charles F. and S.A. Rodgers. "Physic Department Ratings: Another Evaluation," *Science,* No. 174 (1971), pp. 565-68.

Elton, Charles F. and Harriet A. Rose. "What are the Ratings Rating?" *American Psychologist* (March, 1972), pp. 197-201.

Fox, Douglas. "The Identification of Community Leaders by the Reputational and Decision Methods: 3 Cases and Analysis of the Literature," *Sociology and Social Research,* No. 54 (October 1, 1969), pp. 94-103.

Higgins, Stephen. "Quantity and Quality of Doctoral Over-Production," *Educational Record* (Summer, 1971), pp. 262-67.

Jacobson, R.L. "Rating of Graduate Schools Proposed on a Continual Basis," *The Chronicle of Higher Education*, 5, 33 (1971), p. 4.

Marty Martin. "Rating Welch's Ratings." *The Christian Century*, Vol. 89, No. 5 (February 2, 1972), pp. 122-23.

Nichols, J.H. "The Welch Report and the Seminaries," *The Christian Century*, Vol. 89, No. 5 (February 2, 1972), pp. 119-22.

Preston, James O. "The Search for Community Leaders: A Re-Examination of the Reputational Technique," *Sociological Inquiry*, Vol. 39, No. 1 (Winter, 1969), pp. 39-47.

Prochansky, Harold M. "For What Are We Training Our Graduate Students?" *American Psychologist* (March, 1972), pp. 205-12.

Smith, Richard and Fred E. Fiedler. "The Measurement of Scholarly Work: A Critical Review of the Literature," *Educational Record* (Summer, 1971), pp. 225-32.

Tyler, Forrest B. "Knowledgeable Respondents: Private Club or Public Service?" *American Psychologist* (March, 1972), pp. 191-96.

Wispe, L.G. "The Bigger the Better: Productivity, Size, and Turnover in a Sample of Psychology Departments," *American Psychologist*, 24 (1969), pp. 662-68.